




ISSN: 1945-6123

Journal of Registry Management

Spring 2021 | Volume 48 | Number 1

Published by the National Cancer Registrars Association
Founded in 1975 as *The Abstract*



National
Cancer
Registrars
Association

FINDING YOUR DREAM JOB IS JUST A CLICK AWAY

Visit NCRA's Job Bank at www.ncra-usa.org/job!

VIEW JOBS

Access the newest jobs available to cancer registrars seeking employment.

POST AN ANONYMOUS RESUME

Whether you are actively seeking work or curious about what is available, your online resume is your ticket to great job offers!

DESIGN A PERSONAL JOB ALERT

Create job alerts and never let a matching job opportunity pass you by! New jobs that match your search criteria will be e-mailed directly to you.

CREATE A FREE JOB SEEKER ACCOUNT

Your personal job seeker account will allow you to find jobs, manage your resume, and set-up job alerts. Get started today!

NEED A CLINICAL PLACEMENT SITE?

NCRA's Job Bank also includes sites where students can complete their clinical hours and search for mentors.

NCRA'S JOB BANK www.ncra-usa.org/job

**Address change of address and
subscription correspondence to:**

**National Cancer
Registrars Association**
1330 Braddock Place, #520
Alexandria, VA 22314

**Address all editorial
correspondence to:**

Danette Clark, BS, RMA, AAS, CTR,
Editor-in-Chief
Email: JRMEditor@ncra-usa.org
Telephone: (973) 971-5189

Letters to the Editor

Letters to the Editor must be
signed and include address and
telephone number; none will be
published anonymously. Letters
subject to editing.

Editorial Policy

Opinions expressed in articles
appearing in *Journal of Registry
Management* are those of the
authors and not necessarily those
of the National Cancer Registrars
Association, or the editor of
Journal of Registry Management.

**Subscription to *Journal of
Registry Management***

Subscription is a benefit of
membership in the National
Cancer Registrars Association.
Individual subscriptions may
be purchased for \$40/year U.S.
dollars. Single copies may be
purchased for \$12 U.S., \$15
International, prepaid only.

Advertising Sales

Advertising is accepted with the
understanding that products and
services are professionally related
and meet the ethical standards of
public health practice. Acceptance
by the *Journal* does not indicate or
imply endorsement by the *Journal*
or NCRA. Advertising sales
should be directed to
Marketing Coordinator, NCRA,
1330 Braddock Place, #520,
Alexandria, VA 22314,
(703) 299-6640, (703) 299-6620 FAX.

© 2021
**National Cancer
Registrars Association**



Contents

Letter from the Editor

Danette Clark, BS, RMA, AAS, CTR 3

Original Articles

A Case Study of Early-Onset Colorectal Cancer: Using Electronic Health Records
to Support Public Health Surveillance on an Emerging Cancer Control Topic 4

*Julie S. Townsend, MS; Mary Catherine Jones, MPH; Mildred N. Jones, BA, CTR;
Amy W. Waits, BS, CTR; Kamilah Konrad, LMSW; Natasha M. McCoy, MPH*

Exploring Cancer Registrars' Perceptions of the Quality and Use
of Cancer Registry Data for Care Coordination and Other Purposes 12

*Jennifer Peterson, PhD, RHIA, CTR; Shannon H. Houser, PhD, MPH, RHIA, FAHIMA;
Cathy A. Flite, PhD, RHIA, FAHIMA; Lakesha Kinnerson, MPH, RHIA, CPHQ;
Hannah Birchfield, BS, RHIA; April Post, BS*

Late-Stage Diagnosis and Cost of Colorectal Cancer Treatment
in Two State Medicaid Programs 20

*Sonja Hoover, MPP; Sujha Subramanian, PhD; Susan A. Sabatino, MD, MPH;
Jaya S. Khushalani, PhD; Florence K. L. Tangka, PhD*

A Midwest Tri-State Study of Overall Survival in Ovarian Cancer
with Adjuvant Chemotherapy 28

*Sue-Min Lai, PhD, MS, MBA; Li Huang, MPH; Sarma Garimella, MBBS, MPH;
John Keighley, PhD; Michele M. West, PhD*

How I Do It

Attribute Associations as a Practical Approach to Account for Uncertainty
in Residential Geolocation of Cancer Surveillance Data 36

*Christian A. Klaus, MA; Carson Coggins, MS; Kevin A. Henry, PhD; Suzanne Bock, MPH;
Dora Il'yasova, PhD*

Raising the Bar

Information Overload and What You Can Do About It 44

Michele Webb, CTR

Spring 2021 Continuing Education Quiz 46

Cari Vida, BS, RHIA, CTR; Susanna Mitchell, BS, CTR

Call for Papers 48

Information for Authors 49

Editors

Danette A. Clark, BS, RMA, AAS, CTR, Editor-in-Chief
Vonetta Williams, PhD, MPH, CTR, Editor Emeritus
Ginger Clark, Copy Editor
Neisha Johnson, Production Editor

Contributing Editors

Cari Vida, BS, RHIA, CTR—CE Credits
Susanna Mitchell, BS, CTR—CE Credits

2020–2021 NCRA Board of Directors

President: Nadine R. Walker, MS, CTR
President Elect/Secretary: Karen Anne Mason, MSc, RN, CTR
Immediate Past President: Paulette M. Zinkann, CTR
Treasurer Senior: Jennifer H. Garvin, PhD, RHIA, CPHQ, CTR
Treasurer Junior: Leah Kiesow, MBA, CTR
Public Relations & Communications Board Director:
Ronda Broome, MS, CTR
Educational Director: Sharmen A. Dye, CTR
Professional Development Director: Janet L. Vogel, CTR
Recruitment & Retention Director: Ebony Johnson, CTR
ATP Director—East: Vanessa R. Stinson, BSM, CTR
ATP Director—Midwest: Joann Marie Zeller, MBA, RTT, CTR
ATP Director—West: Kelli K. Olsen, MS, CTR

NCRA Staff

Neisha Johnson, Membership Manager
Peggy Meehan, Director, Education and Communications

2020–2021 JRM Advisory Board

Baozhen Qiao, MD, PhD
Roneé E. Wilson, PhD, MPH
Tammy Berryhill, MA, RHIA, CTR
Monique Hernandez, PhD
Donna M. Gress, RHIT, CTR
Amy Kahn, MS, CTR
Danillie Clark, RMA, AAS, CTR
Pamela Webster, PhD, CTR
Mary Elizabeth O'Neil, MPH
Jason L. Salemi, PhD, MPH
April A. Austin, MS
Andrea Sipin, MS, CPHI, CTR
Sarah Nash, PhD, MPH, CPH
Keri Miller, MHA CTR
Vonetta Williams, PhD, MPH, CTR (Editor Emeritus)
Reda Wilson, MPH, CTR (Editor Emeritus)

Indexing

The *Journal of Registry Management* is indexed in the National Library of Medicine's MEDLINE database. Citations from the articles indexed, the indexing terms (key words), and the English abstract printed in *JRM* are included and searchable using PubMed.

For your convenience, the *Journal of Registry Management* is indexed in the 4th issue of each year and on the Web (under "Resources" at <http://www.ncra-usa.org/jrm>). The 4th issue indexes all articles for that particular year. The Web index is a cumulative index of all *JRM* articles ever published.

Dear Colleagues,

It's been a full year of overcoming the challenges of the COVID-19 pandemic. Registries have adapted to working remotely and virtual meetings are now the new normal. The wonderful season of spring has arrived. The theme for the 47th Annual National Cancer Registrars Conference is "Driving Data Into The Future." The conference will be virtual again this year and covers a wide variety of topics. Another opportunity to expand your knowledge comes from the 2021 SEER Advanced Topics for Registry Professionals Workshop. I hope to see you virtually at both events.

We have a nice collection of original articles in this edition, starting with Julie S. Townsend, MS, and colleagues, who discuss a case study using electronic health records for surveillance for certain types of cancer. Jennifer Peterson, PhD, RHIA, CTR, and team explore the use of registry data for care coordination. Next, Sonja Hoover, MPP, and associates review the cost of colorectal cancer treatment in Medicaid programs. The final original article is from Sue-Min Lai, PhD, MS, MBA, and coauthors, who studied survival in ovarian cancer with adjuvant chemotherapy.

In our "How I Do It" section, Christian A. Klaus, MA, and associates cover the topic of geolocation of cancer surveillance data. The final item is our "Raising the Bar" feature with Michele Webb's take on information overload.

I have a request for the readership. The last year has been difficult for facilities who struggled to meet the challenges of the pandemic. I know there are registries who made changes that others can benefit from. Please consider sharing your successes with us. Topics may include staffing, tumor conferences, and cancer committee. This is not just limited to hospital registries. Central registries had to make changes as well during the pandemic. These items would be included in our "How I Do It" section.

The last 2 pages of the JRM contains the *Call for Papers* and *Information for Authors*. Submissions of manuscripts or articles are accepted at any time. The "How I Do It" section comes from readers who want to share their expertise and ideas on varying topics. A few requested topics are case-finding, class of case, follow-up, and COVID-19 data items.

Regards,

Danette A. Clark, BS, RMA, AAS, CTR
JRMeditor@NCRA-USA.org

A Case Study of Early-Onset Colorectal Cancer: Using Electronic Health Records to Support Public Health Surveillance on an Emerging Cancer Control Topic

Julie S. Townsend, MS^a; Mary Catherine Jones, MPH^b; Mildred N. Jones, BA, CTR^c; Amy W. Waits, BS, CTR^c; Kamilah Konrad, LMSW^c; Natasha M. McCoy, MPH^b

Abstract: Electronic health records (EHRs) are increasingly being used to support public health surveillance, including in cancer, where many population-based registries can now accept electronic case reporting. Using EHRs to supplement cancer registry data provides the opportunity to examine in more detail emerging issues in cancer control, such as the increasing incidence rates of early onset colorectal cancer (CRC). The purpose of this study was to evaluate the feasibility of a public health organization partnering with a health system to examine risk factors for early-onset CRC in a community cancer setting, and to further understand challenges with using EHRs to address emerging topics in cancer control. We conducted a mixed-methods evaluation using key informant interviews with public health practitioners, researchers, and registry staff to generate insights on how using EHRs and partnering with health systems can improve chronic disease surveillance and cancer control. A data quality assessment of variables representing risk factors for CRC and other clinical characteristics was conducted on all CRC patients diagnosed in 2016 at the participating cancer center. The quantitative assessment of the EHR data revealed that, while most chronic health conditions were well documented, around 25% of CRC patients were missing information on body mass index, alcohol, and tobacco use. Key informants offered ideas and ways to overcome challenges with using EHR data to support chronic disease surveillance. Their recommendations included the following activities: engaging EHR vendors in the development of standards, taking leadership roles on workgroups to address emerging technological issues, participating in pilot studies and task forces, and negotiating with EHR vendors so that clinical decision support tools built to support public health initiatives are freely available to all users of those EHRs. Although using EHR data to support public health efforts is not without its challenges, it soon could be an important part of chronic disease surveillance and cancer control.

Key words: colorectal neoplasms, electronic health records, public health surveillance, registries

Introduction

Electronic health records (EHRs) are increasingly being used to support public health surveillance for a variety of health topics, including syndromic surveillance, immunizations, cancer, childhood obesity, diabetes, hypertension, asthma, as well as other conditions.¹⁻⁵ While electronic laboratory records (ELRs) have supported public health surveillance for a number of years, EHRs are also being tapped to enhance reporting of immunizations and notifiable conditions to health departments and registries, including infectious diseases, and incident cancer cases (<https://www.naaccr.org/meaningful-use/>).^{1,6-8}

While incidence rates are declining for many cancer sites, there are notable increases in rates of new cases of early-onset colorectal cancer (CRC); uterine, pancreatic, and liver cancers; as well as others in which there is a need to better understand contributing factors to rising

incidence rates and implement evidence-based interventions to address them.^{9,10} Early-onset CRC was also chosen as a topic because of the importance of genomic testing in this disease and the opportunities available for identifying at-risk family members through cascade testing for Lynch syndrome and other inherited disorders. Additionally, understanding the prevalence of obesity as well as alcohol and tobacco use and other chronic health conditions may be important to cancer control planners, given that early-onset CRC survivors may benefit from survivorship care interventions to mitigate adverse health outcomes and improve their quality of life.¹¹

Population-based central cancer registries are the backbone of cancer surveillance in the United States, and provide valuable information on demographic and clinical characteristics of new cancer cases (<https://www.cdc.gov/uscs>). However, there are opportunities for EHR, laboratory, and

^aDivision of Cancer Prevention and Control, Centers for Disease Control and Prevention, Atlanta, Georgia. ^bNational Association of Chronic Disease Directors, Decatur, Georgia. ^cNorthside Hospital Cancer Institute, Atlanta, Georgia.

Address correspondence to Julie S. Townsend, MS, Centers for Disease Control and Prevention, 4770 Buford Hwy, NE, MS S107-4, Chamblee, GA 30341. Email: jtownsend@cdc.gov. Telephone: (770) 488-3012.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the National Association of Chronic Disease Directors. Funding for the Health Behaviors and Other Risk Factors in Early-Onset Colorectal Cancer project was provided by Cooperative Agreement Number 5NU38OT000225-05, sponsored by the Centers for Disease Control and Prevention, Division of Cancer Prevention and Control (CDC), and the National Association of Chronic Disease Directors (NACDD).

This project was considered routine public health practice and did not require Institutional Review Board approval.

health systems data to support or enhance data reported to cancer registries and provide information not commonly found in registries (eg, tobacco use or obesity) to support special surveillance studies for emerging public health challenges.¹²⁻¹⁵ Some cancer registries have been linked to administrative and claims data and other sources.¹⁵⁻¹⁸ In many research settings, “big data” is increasingly being used to support cancer research studies, and cloud-based infrastructure has brought together hospital cancer registry, EHR, laboratory, and pharmacy data to drive improvements in cancer treatment and care.^{19,20} Linking to EHRs and other electronic data sources represents a new opportunity to examine cancer risk factors, screening test use, molecular characteristics, and chronic health conditions (among others) before or at the time of a cancer diagnosis, as well as serving as another source of information on treatment, treatment adverse effects, recurrence, and other health conditions for public health researchers and practitioners.²¹⁻²⁵

In 2017–2019, we undertook a pilot project to assess the feasibility of partnering with a large health system (Northside Hospital) in Atlanta, Georgia to assess risk factors for early-onset CRC and opportunities for prevention and control among the patient population at its community cancer center (Northside Hospital Cancer Institute), using both EHR and cancer registry data. As part of the project’s evaluation, we completed a data quality assessment of EHR variables and conducted key informant interviews with public health professionals and cancer registry staff to identify successes, challenges, and barriers to using EHR and other health systems data to support chronic disease surveillance and special investigations, with the goal of identifying recommendations for public health departments that are interested in partnering with health systems on cancer-related projects at the local level to inform comprehensive cancer control efforts.

Methods

Project Description

This project was a joint collaboration between the National Association of Chronic Disease Directors, the Centers for Disease Control and Prevention (CDC), and the Northside Hospital Cancer Institute, an American College of Surgeons Commission on Cancer (CoC)-accredited comprehensive community cancer center with 3 acute-care hospitals serving the metropolitan Atlanta area at the time of the study. Briefly, one of the primary activities of our pilot project was conducting a descriptive, retrospective study of all 2016 CRC cancer cases at Northside Hospital Cancer Institute. Data elements were selected based on known CRC risk factors from the scientific literature, available registry data, and emerging conditions of interest. We obtained permission so that our certified tumor registrars (CTRs) could access and abstract data from 2 EHR systems: the hospital’s and an affiliated gastroenterology practice. CTR team members went through an initial 10-case quality assurance (QA) review for each abstractor, with feedback and education provided along with an ongoing 10% QA review by a senior CTR team member to ensure data completeness and accuracy.

Mixed-Methods Evaluation Approach

Given that the project was a feasibility study, an evaluation plan using a mixed-methods approach to collect both quantitative and qualitative data was developed to address the specific study questions of the project; namely:

1. Can medical data from an integrated health care delivery system be rapidly assessed and used to determine accurate and high-quality information on early onset CRC without the need to contact the patient?
2. How can we use what we learn to build capacity among other integrated health care delivery systems and their public health partners, particularly those in the community cancer setting?

Quantitative Data Quality Assessment

We addressed the first question through an analysis conducted during a data-quality assessment of key variables needed to assess potential risk factors for early-onset CRC not typically collected as part of the cancer registry abstract. These included body mass index (BMI), tobacco use, alcohol use, CRC screening history, tumor screening for Lynch syndrome, and family history of CRC and related Lynch syndrome cancers (Figure 1). Other variables collected and assessed included demographic characteristics (eg, driving distance from the patient’s residence to the cancer center, patient’s preferred language, and patient’s status as a caregiver). Clinical characteristics included a history of chronic health conditions, such as inflammatory bowel disease and diabetes. In total, an additional 114 data elements were abstracted from 2 EHR systems. During the data-quality assessment, we analyzed the number of patients with missing or unknown information using SAS statistical software.

Qualitative Key Informant Interviews

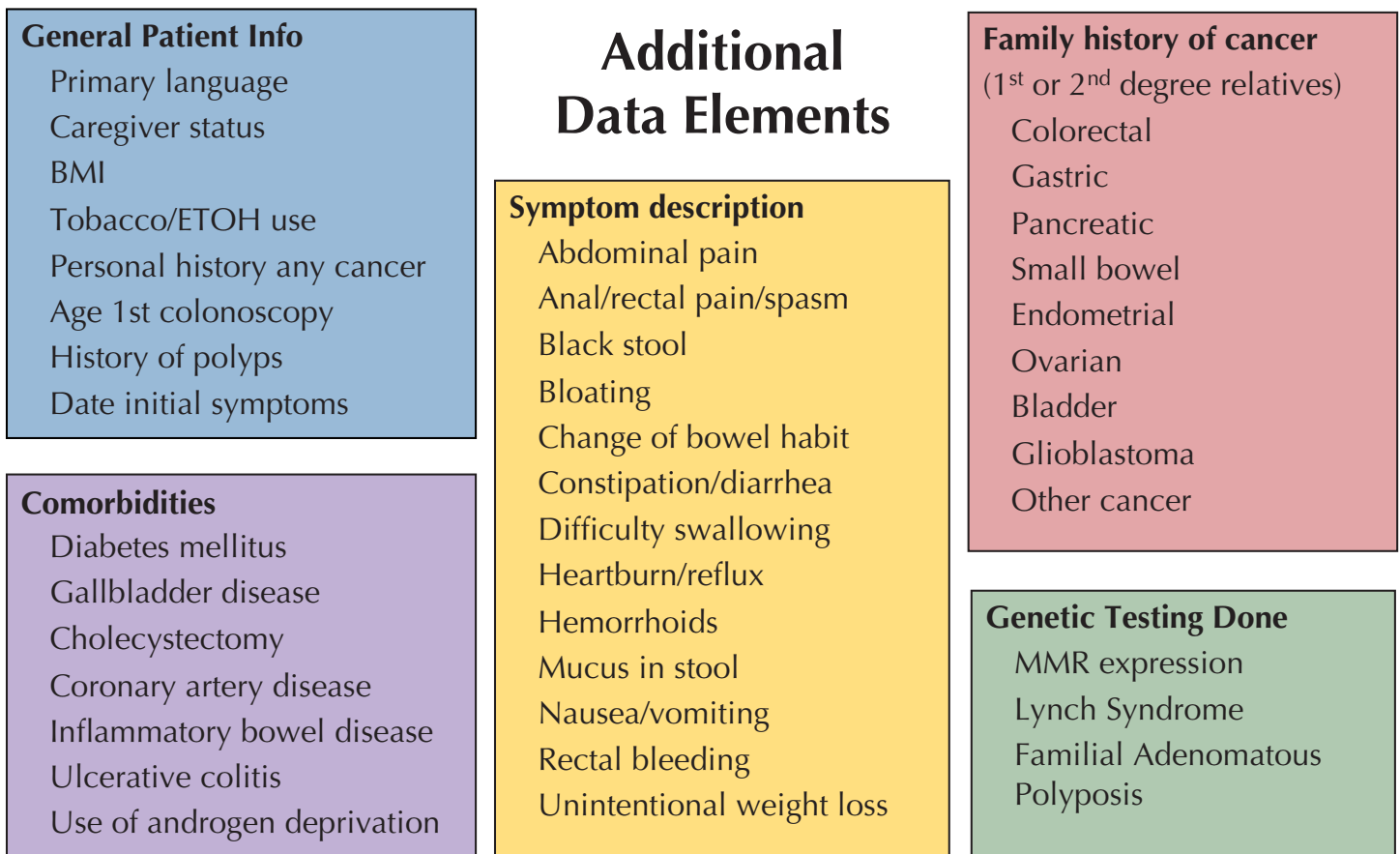
We completed qualitative key informant interviews with subject matter experts in the areas of laboratory reporting, state cancer registries, hospital cancer registries, state-level chronic disease epidemiology, and syndromic surveillance to better understand how to build capacity among other integrated health care delivery systems and their public health partners. Interviewees were selected based on project team recommendations with the goal of including different professional experiences with EHRs from the public health field. With the exception of 1 expert who did not respond to our inquiry, all experts invited to an interview completed one.

A semistructured discussion guide was developed that addressed the following:

1. The current landscape of using health systems data and EHRs to support public health surveillance
2. The facilitators (ie, keys to success) and barriers to health care systems partnering with public health organizations on surveillance efforts
3. The processes, policies, or practices that can help to overcome the barriers and capitalize on facilitating factors

Interviews were conducted with 9 subject matter experts who had experience with EHRs and health systems

Figure 1. Select Additional Data Elements Collected for the Colorectal Cancer Cohort



data to support public health surveillance or research. Subject matter experts included cancer center registry staff, state cancer registry staff, a gastroenterologist, and employees at state and federal government agencies. All interviews were conducted via telephone, except for 1 in-person team interview with cancer center staff involved in the project. Discussions were conducted in segments of 30 to 60 minutes. The team evaluator led key informant interviews and involved team members in contributing to the discussion with subject matter experts, including providing contextual information about the project, encouraging authentic discussion, and asking follow-up questions to prompt for additional insights and observations. The team evaluator took notes during the sessions and synthesized themes that emerged through the discussions with the subject matter experts. The themes were organized according to the potential audience (public health professionals, health systems, and industry/professional organizations) and recommendations/actions that could be done by the audience to advance the use of EHRs for public health surveillance.

Study Approval

CDC review determined this project to be public health practice. Office of management and budget approval was not required for data collection because fewer than 9 nonfederal key informants were interviewed, and information was collected through secondary data sources for the

data assessment. The data assessment was approved by the Northside Hospital Research Oversight Committee.

Results

What are Factors that Influence the Quality and Accuracy of, and Ease of Access to, Information on Early-Onset Colorectal Cancer Risk Factors in the Cancer Center's Records?

Findings from the data quality assessment revealed that data completeness (percentage of unknown or missing information) varied, depending on the variable collected (Table 1). Unknown/missing values ranged from 5% for common chronic health conditions to around 25% for health behaviors like alcohol and tobacco use. Around 25% of CRC patients had missing information on their BMI, and similar proportions had missing information for a family history of CRC or endometrial or ovarian cancer. Variables with the highest percentage of missing data or unknown information included the patient's preferred language (27%), the patient's caregiving status (36%), history of polyps (31%), history of a previous cancer (31%), and time from onset of symptoms to diagnosis (33%).

Key informants from the cancer center who participated in this project worked with multiple practices that use different EHRs, which did not necessarily "communicate" with each other. The use of multiple EHRs complicated data analysis for various reasons, including variation among

Table 1. Findings from Data Quality Assessment of Key Variables for CRC Cohort, N = 721

Variable	Unknown/ missing, n (%)
Demographic characteristics	
Distance to hospital from residence	55 (7.6)
Primary language spoken	191 (26.5)
Patient caregiver status	257 (35.6)
Health behaviors	
Alcohol use history	178 (24.7)
Tobacco use history	160 (22.2)
Body mass index	186 (25.8)
Clinical factors	
Time from initial symptom to diagnosis	240 (33.3)
History of other cancer	220 (30.5)
History of polyps	222 (30.8)
Diabetes mellitus	37 (5.1)
Gallbladder disease	37 (5.1)
Coronary artery disease	37 (5.1)
Inflammatory bowel disease	37 (5.1)
Family history of cancer	
First-degree relative with CRC	180 (25.0)
First-degree relative with endometrial cancer	176 (24.4)
First-degree relative with ovarian cancer	175 (24.3)
Second-degree relative with CRC	176 (24.4)
Second-degree relative with endometrial cancer	175 (24.3)
Second-degree relative with ovarian cancer	175 (24.3)

CRC, colorectal cancer.

EHRs in the headings, fields, and ways that risk factor data are documented, which challenged analysis across EHRs. There were also different security requirements for each EHR. Within EHRs, there was inconsistency in how and where information is documented. Some data were found in multiple locations within the EHR, such as patient history and the intake form, and sometimes the information conflicted. A physician's office may collect the information differently from a surgical preadmission form about the same topic (eg, do you smoke vs history of smoking/ever smoked). Key informants noted that there needs to be a protocol for determining which data to consider for risk factors. Additionally, key informants noted inconsistencies among providers in how often risk factor data were updated. For example, family history may be collected at intake but never updated over the course of the patient's care.

Some medical information continues to be collected on paper, outside of EHRs. We collected some data elements needed for the analysis from documents scanned into the EHR rather than entered into electronic fields. This had to be retrieved manually, which slowed the assessment and

added cost in staff time. We learned that extending the use of the EHR from patient care to surveillance requires a shift in how the data are collected and analyzed. Creating user-defined fields in the cancer registry software to capture information not readily available in existing fields was resource-intensive.

What are Systemic Factors that Influence the Availability and Quality of Electronic Health Data That Can be Used for Public Health Surveillance, Especially for Emerging Issues in Cancer Control?

Subject matter experts noted that the implementation of the American Reinvestment and Recovery Act—Health Information Technology for Economic and Clinical Health (HITECH), which spurred adoption of EHRs by hospitals and physician practices, and Meaningful Use, which laid additional groundwork for health systems' data to be used for public health, have substantially increased the availability and quality of EHR data. Various national-level mechanisms to promote quality improvement and value-based care, including reporting on quality standards, have been instrumental. The CDC has worked to engage health systems, EHR vendors, and other stakeholders in surveillance of cancer, immunizations, asthma, diabetes, and syndromes. Key informants noted that a competitive marketplace, including competition among EHR vendors for market share and competition among health care providers for patients may have influenced the availability and quality of EHR data, the increased consumption of health care quality information by patients, and the increased use of technology for patients to monitor and report health care data to providers. Additionally, professional organizations have developed support to health systems staff in improving quality and accuracy of patient information in cancer registries.

Despite the convergence of these factors in increasing the availability and use of EHR data, multiple key informants noted the workload challenges that the technology presents to health care providers. One key informant noted that, despite advancements in EHR technology, they have not necessarily made data collection and entry more efficient or translated into more time for patient care:

"Finding a way to make up for the increased workflow required by data entry is a challenge. The way EHR systems are constructed is very old school, the electronic version of someone taking notes or writing them in a paper record."

What are Some Opportunities for Health Systems, Public Health, and Allied Agencies to Increase the Value and Use of Electronic Health Record Data for Chronic Disease Surveillance, Especially the Identification of Risk Factors for Early-Onset Colorectal Cancer?

Table 2 summarizes recommendations for what health systems, public health agencies, and professional and industry associations can do to improve the use of EHR data for chronic disease surveillance, organized along with the themes that emerged from the qualitative data analysis of the interviews. These themes included engaging stakeholders/

Table 2. Thoughts and Ideas Provided by Key Informants for Improving Electronic Health Records to Support Chronic Disease Surveillance and Cancer Control, by Audience Type and Topic Area

<i>Health Systems</i>	<i>Public Health</i>	<i>Professional and Industry Associations</i>
Building partnerships/engaging stakeholders		
<p>Have conversations with public health organizations about their common interests in chronic disease prevention, including screening and early identification, and management.</p> <p>Bring together providers and EHR vendors to adapt EHRs to provide the greatest benefit to practice and administration.</p> <p>Engage providers in developing the clinical questions that align with the public health surveillance needs.</p> <p>Have conversations with public health organizations about their common interests in chronic disease prevention, including screening and early identification, and management.</p>	<p>Coordinate across disease areas and administrative units to work with health systems, including building on existing efforts related to syndromic surveillance and immunization where appropriate.</p> <p>Collaborate across different levels of government to ensure that local, state, and federal opportunities are connected.</p> <p>Facilitate the connection of health systems with community organizations, public health coalitions, and other stakeholders that can use health systems data to inform community health and action planning.</p> <p>Collaborate with health systems to identify priority health topics or disease areas where the health system and community may derive significant benefit from EHR improvement.</p> <p>Explore opportunities to work with health plans to incentivize identification of early onset CRC risk factors for providers.</p>	<p>Convene diverse stakeholders at different levels of healthcare organizations, including members, with public health to facilitate problem definition and strategic planning.</p> <p>Learn the landscape of health information technology and its intersections with public health and assist other organizations in defining their roles.</p> <p>Engage EHR vendors in collaborations, especially in the development of standards and discussions of how to increase interoperability.</p>
Taskforce/workgroup participation		
<p>Participate in state or national level pilot studies, task forces, and workgroups to identify and implement mutually beneficial opportunities for collaboration with public health.</p>	<p>Lead or participate in the development of standards for EHR vendors.</p>	<p>Lead regional or national workgroups and task forces to address collaboration between public health and health systems around surveillance, such as the CDC/CSTE/APHL's Electronic Laboratory Reporting Task Force, and the CDC Public Health-EHR Vendors Collaboration Initiative.</p>
Administrative/system changes		
<p>Encourage or incentivize their affiliated practices to use one EHR, which may eliminate inefficiencies for data collection and analysis associated with the use of multiple EHR platforms.</p>	<p>Assist health systems in identifying where quality improvement and financial incentives align with public health surveillance priorities.</p> <p>Create and promote recognition programs to highlight the efforts of health systems making strides in using their EHR for chronic disease surveillance.</p>	<p>Provide incentives to pilot novel approaches to EHR vendor and health systems collaborations to support chronic disease surveillance; obtain agreement from the vendor that any products resulting from the collaboration are available freely to all providers using that EHR system.</p>
Improving data quality and data use		
<p>Develop, implement, train on, and enforce protocols and standards for documentation among providers.</p>	<p>Encourage the use of existing public health data sources and healthcare data in community health needs assessments.</p> <p>Demonstrate the connection between public health surveillance interests and clinical quality measures.</p> <p>Engage EHR vendors in piloting algorithms that can improve the ease of collecting early onset CRC risk factor data.</p>	<p>Continue the development of tools and resources that health systems and public health can use or adapt to facilitate EHR modification to meet public health surveillance interests, such as the Healthcare Information and Management Systems Society's (HIMSS) EMR Usability Evaluation Toolkit.³⁹</p>
Communication/dissemination		
<p>Communicate the direct benefit of improved documentation to providers, including sharing of findings from surveillance efforts and how improved documentation links back to improved patient care.</p>	<p>Share successes and challenges and develop solutions with other public health organizations.</p>	<p>Promote the exchange of successful public health-health systems data partnerships within and across disease areas through publications, conference presentations, and webinars.</p> <p>Define the mutual benefit of public health-health systems collaborations on surveillance.</p>

building partnerships, task force/work group participation, administrative/systems change, improving data quality and data use, and communication/dissemination. Notably, health systems and public health agencies can collaborate to improve population health, and one starting point is participating in state or national level pilot studies, task forces, and work groups. There are opportunities for health systems to engage providers in developing clinical decision support tools that are implemented with population health management platforms with existing interfaces with EHRs. Health systems can also develop and enforce protocols and standards for documentation among providers.

Public health entities may consider assisting health systems in identifying where quality improvement and financial incentives align with public health surveillance priorities as a way to build support and generate interest in using EHR data for chronic disease surveillance. At the local level, public health agencies could collaborate with health systems to identify priority health topics or disease areas where the health system and community would derive significant benefit from EHR improvement (eg, diabetes management). Public health agencies could assist health systems in identifying shared data needs related to surveillance and with vetting variables to be monitored and reported. Within public health agencies, coordination across disease areas and administrative units to work with health systems, including building on existing efforts related to syndromic surveillance and immunization, may synergize efforts and increase efficiencies for chronic disease surveillance.

Allied organizations, including professional and industry associations, could engage EHR vendors in collaborations, especially in developing standards and discussions of how to increase interoperability. They could also provide incentives to pilot novel approaches to EHR vendor and health systems' collaborations to support chronic disease surveillance and obtain agreements from vendors that any products resulting from the collaboration are available freely to all providers using that EHR system.

Discussion

The interviews with key informants on using EHRs and health systems data to support the investigation of emerging topics in cancer control and other chronic disease surveillance activities revealed several key domains: forming partnerships/engaging stakeholders, participating in task forces/work groups, providing education on systems and administrative changes, improving EHR data quality, and communicating/disseminating findings. Addressing each of these domains may improve the use of EHRs to support cancer control and other public health efforts at the local level. Public health, health systems, and professional/industry associations can all play a role across these domains.

We specifically addressed one key theme, improving EHR data quality, through our own data quality assessment. Our health system partner's EHR and those of its affiliated gastroenterology practices could readily provide information on the prevalence of chronic health conditions,

but around one-fourth of CRC patients each had incomplete data for family history of cancer, health behaviors, and other clinical factors (duration of symptoms, history of polyps, tumor testing, etc) that could be of interest to public health partners who need relevant data to inform community interventions. Although some patients with missing data may have come from nonaffiliated gastroenterology practices outside of the health system, our findings on data completeness may be typical with EHR data.^{26,27} Even well-established health care research networks using virtual data warehouses have noted the challenges with electronic capture of molecular data, particularly data elements that may only be found in scanned imaging reports and are not captured in standardized EHR data fields or as site-specific factors in tumor registries.²⁵ Even with these limitations, analysis of available data may be helpful for public health surveillance purposes and generating new hypotheses, which can be tested further in prospective studies. Additionally, it provides a snapshot of care that goes beyond analyzing traditional cancer registry data elements, which can provide helpful local data that public health partners could potentially use to improve access to care and train providers on use of clinical guidelines, such as tumor screening for Lynch syndrome and genetic counseling referral.²⁸

Although EHRs are increasingly being used to support chronic disease surveillance, their use so far has been limited to a few topic areas (eg, diabetes, obesity, asthma, hypertension, cancer electronic reporting) and geographic areas of the United States. However, partnerships between public health organizations and health systems are increasingly becoming common to address a variety of chronic health conditions and implement interventions to improve health. For example, CDC-funded cancer programs at health departments and universities are partnering with health systems on projects to increase cancer screening and generate survivorship care plans.^{11,29} Given that these efforts rely on accurate data, projects such as these may contribute to overall improvements in useful and quality data.

Health care technology is constantly evolving, and it may be challenging for public health organizations to keep up with new technologies, like HL7 Fast Healthcare Interoperability Resources (<https://www.hl7.org/fhir/overview.html>), that can streamline data exchange and make it possible to get regular data feeds so that the most current patient data are available.³⁰ Other technologies are increasingly becoming available that may help standardize data across different EHR platforms and capitalize on natural language processing techniques to make data more accessible for public health needs.^{31,32} As distributed data networks become more commonplace, it may be increasingly important for public health organizations to be engaged with health systems around data so that emerging topics in cancer control can be quickly assessed and appropriate interventions and timely access to clinical care applied.³³⁻³⁵ These efforts hinge on having public health and health system partners with adequate skills in data science and the information technology infrastructure for big data.³⁶ CTRs may continue to play a key role in ensuring

data quality, along with their expertise in the types of information to capture and consolidation of data across multiple information streams.³⁷

There are some limitations to our mixed-methods evaluation. During the EHR data quality assessment, we did not evaluate internal validity, whether certain patient or provider characteristics and referral patterns played a role in data completeness, or if the sample with complete information was representative of the overall patient population. Therefore, we did not evaluate all potential domains of data quality proposed for assessing EHR data for research use.³⁸ We only examined 1 year of data for one cancer site (CRC) at a single health system, limiting generalizability to other patient populations. Although our key informants represented federal and state health departments and health systems perspectives, we did not recruit key informants employed by EHR vendors, who may have lent a different perspective on EHR use.

Despite these limitations, there are some strengths to our study. We evaluated a practical use case scenario using cancer registry data supplemented by EHR data elements to better understand risk factors among early- and late-onset CRC patients. We were able to leverage trained CTRs using a data dictionary that we developed to capture standardized information from divergent EHRs. Our key informants represented many different user experiences and lent valuable insights into using EHRs to support investigating emerging topics in cancer control.

Conclusion

Major efforts are underway at the federal, state, academic, and local health care levels to tap into EHRs, laboratory data, biobanks, and genomics data to integrate information for a more complete picture of population health.^{6,26} The key domains we identified through our key informant interviews may be able to guide public health practitioners, health systems, and professional associations/vendors on how to navigate this uncharted territory by providing concrete actions that can be undertaken through this journey. Our data quality findings may be used to identify problem areas in EHRs that need attention, such as improving the documentation of health behaviors and cancer family history that may impact the cancer patient's prognosis through the treatment and survivorship period, demographic characteristics related to the social determinants of health, and other clinical characteristics that can inform community-level interventions with health system partners. Big data analytics using integrated, cloud-based data may one day allow public health professionals, researchers, and cancer control planners to better understand emerging topics in cancer control, including early-onset CRC.

Acknowledgements

We would like to thank our project team members: Maurshanda Renee Matthews, Sheema Ahmed, Dr. Marc Sonenshine, Kate Canterbury, Dr. Xu Xiang, Dr. C. Brooke Steele, Frank Bright, and Dr. Paulette Valliere for their contributions to this project. We would like to acknowledge

the key informants from the CDC, National Cancer Institute, University of Kentucky, and Hawaii Department of Health who kindly agreed to provide their thoughts and opinions about using electronic health records to support chronic disease surveillance and cancer control.

References

1. Birkhead GS, Klompas M, Shah NR. Uses of electronic health records for public health surveillance to advance public health. *Annu Rev Public Health*. 2015;36:345-359.
2. Perlman SE, McVeigh KH, Thorpe LE, Jacobson L, Greene CM, Gwynn RC. Innovations in population health surveillance: using electronic health records for chronic disease surveillance. *Am J Public Health*. 2017;107(6):853-857.
3. Gutilla MJ, Davidson AJ, Daley MF, Anderson GB, Marshall JA, Magzamen S. Data for community health assessment in rural Colorado: a comparison of electronic health records to public health surveys to describe childhood obesity. *J Public Health Manag Pract*. 2017;23(suppl 4):S53-S62.
4. Klompas M, Cocoros NM, Menchaca JT, et al. State and local chronic disease surveillance using electronic health record systems. *Am J Public Health*. 2017;107(9):1406-1412.
5. Thorpe LE, McVeigh KH, Perlman S, et al. Monitoring prevalence, treatment, and control of metabolic conditions in New York City adults using 2013 primary care electronic health records: a surveillance validation study. *EGEMS (Wash DC)*. 2016;4(1):1266.
6. Cooney MA, Iademarco MF, Huang M, MacKenzie WR, Davidson AJ. The public health community platform, electronic case reporting, and the digital bridge. *J Public Health Manag Pract*. 2018;24(2):185-189.
7. Blumenthal W, Alimi TO, Jones SF, et al. Using informatics to improve cancer surveillance. *J Am Med Inform Assoc*. 2020;27(9):1488-1495.
8. Pollack LA, Jones SF, Blumenthal W, et al. Population health informatics can advance interoperability: National Program of Cancer Registries Electronic Pathology Reporting Project. 2020;4:985-992.
9. Siegel RL, Fedewa SA, Anderson WF, et al. Colorectal cancer incidence patterns in the United States, 1974–2013. *J Natl Cancer Inst*. 2017;109(8):djw32.
10. Henley SJ, Ward EM, Scott S, et al. Annual report to the nation on the status of cancer, part I: national cancer statistics. *Cancer*. 2020;126(10):2225-2249.
11. Rohan EA, Miller N, Bonner F III, et al. Comprehensive cancer control: promoting survivor health and wellness. *Cancer Causes Control*. 2018;29(12):1277-1285.
12. Hernandez MN, Voti L, Feldman JD, et al. Cancer registry enrichment via linkage with hospital-based electronic medical records: a pilot investigation. *J Registry Manag*. 2013;40(1):40-47.
13. Ruppert LP, He J, Martin J, et al. Linkage of Indiana State Cancer Registry and Indiana Network for Patient Care Data. *J Registry Manag*. 2016;43(4):174-178.
14. Reams C, Powell M, Edwards R. State synergies and disease surveillance: creating an electronic health data communication model for cancer reporting and comparative effectiveness research in Kentucky. *EGEMS (Wash DC)*. 2014;2(2):1064.
15. Tucker TC, Durbin EB, McDowell JK, Huang B. Unlocking the potential of population-based cancer registries. *Cancer*. 2019;125(21):3729-3737.
16. Gallaway MS, Huang B, Chen Q, et al. Identifying smoking status and smoking cessation using a data linkage between the Kentucky Cancer Registry and health claims data. *JCO Clin Cancer Inform*. 2019;3:1-8.
17. Blanchette PS, Chung H, Pritchard KI, et al. Influenza vaccine effectiveness among patients with cancer: a population-based study using health administrative and laboratory testing data from Ontario, Canada. *J Clin Oncol*. 2019;37(30):2795-2804.
18. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care*. 2002;40(8 suppl):IV-3-18.
19. Rubinstein SM, Warner JL. CancerLinQ: origins, implementation, and future directions. *JCO Clin Cancer Inform*. 2018;2:1-7.

20. Tran Q, Warren JL, Barrett MJ, et al. An evaluation of the utility of big data to supplement cancer treatment information: linkage between IQVIA pharmacy database and the Surveillance, Epidemiology, and End Results Program. *J Natl Cancer Inst Monogr.* 2020;2020(55):72-81.
21. Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural language processing of clinical notes on chronic diseases: systematic review. *JMIR Med Inform.* 2019;7(2):e12239.
22. Banerjee I, Bozkurt S, Caswell-Jin JL, Kurian AW, Rubin DL. Natural language processing approaches to detect the timeline of metastatic recurrence of breast cancer. *JCO Clin Cancer Inform.* 2019;3:1-12.
23. Dewdney SB, Lachance J. Electronic records, registries, and the development of "big data": crowd-sourcing quality toward knowledge. *Front Oncol.* 2016;6:268.
24. Booth CM, Karim S, Mackillop WJ. Real-world data: towards achieving the achievable in cancer care. *Nat Rev Clin Oncol.* 2019;16(5):312-325.
25. Burnett-Hartman AN, Udaltsova N, Kushi LH, et al. Clinical molecular marker testing data capture to promote precision medicine research within the cancer research network. *JCO Clin Cancer Inform.* 2019;3:1-10.
26. Hiatt RA, Tai CG, Blayney DW, et al. Leveraging state cancer registries to measure and improve the quality of cancer care: a potential strategy for California and beyond. *J Natl Cancer Inst.* 2015;107(5):djv047.
27. Wright A, McCoy AB, Hickman TT, et al. Problem list completeness in electronic health records: a multi-site study and assessment of success factors. *Int J Med Inform.* 2015;84(10):784-790.
28. Provenzale D, Gupta S, Ahnen DJ, et al. Genetic/familial high-risk assessment: colorectal version 1.2016, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw.* 2016;14(8):1010-1030.
29. Tangka FKL, Subramanian S, Hoover S, et al. Identifying optimal approaches to scale up colorectal cancer screening: an overview of the centers for disease control and prevention (CDC)'s learning laboratory. *Cancer Causes Control.* 2019;30(2):169-175.
30. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc.* 2016;23(5):899-908.
31. Mandl KD, Kohane IS. A 21st-Century health IT system: creating a real-world information economy. *N Engl J Med.* 2017;376(20):1905-1907.
32. Warner JL, Levy MA, Neuss MN, Warner JL, Levy MA, Neuss MN. ReCAP: Feasibility and accuracy of extracting cancer stage information from narrative electronic health record data. *J Oncol Pract.* 2016;12(2):157-158; e169-157.
33. Tabano DC, Cole E, Holve E, Davidson AJ. Distributed data networks that support public health information needs. *J Public Health Manag Pract.* 2017;23(6):674-683.
34. Khan S, Shea CM, Qudsi HK. Barriers to local public health chronic disease surveillance through health information exchange: a capacity assessment of health departments in the Health Information Network of South Texas. *J Public Health Manag Pract.* 2017;23(3):e10-e17.
35. Bacon E, Budney G, Bondy J, et al. Developing a regional distributed data network for surveillance of chronic health conditions: the Colorado Health Observation Regional Data Service. *J Public Health Manag Pract.* 2019;25(5):498-507.
36. McFarlane TD, Dixon BE, Grannis SJ, Gibson PJ. Public health informatics in local and state health agencies: an update from the Public Health Workforce Interests and Needs Survey. *J Public Health Manag Pract.* 2019;25(2 suppl):S67-S77.
37. Peterson J. Encouraging health information management graduates to pursue cancer registry careers. *J Registry Manag.* 2016;43(1):37-41.
38. Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. *EGEMS (Wash DC).* 2017;5(1):14.
39. Nine essential principles of software usability. Healthcare Information and Management Systems Society website. Published April 9, 2015. Accessed December 4, 2019. <https://www.himss.org/himss-emr-usability-evaluation-guide-clinicians-practices-9-essential-principles-software-usability>

Exploring Cancer Registrars' Perceptions of the Quality and Use of Cancer Registry Data for Care Coordination and Other Purposes

Jennifer Peterson, PhD, RHIA, CTR^b; Shannon H. Houser, PhD, MPH, RHIA, FAHIMA^b; Cathy A. Flite, PhD, RHIA, FAHIMA^c; Lakesha Kinnerson, MPH, RHIA, CPHQ^d; Hannah Birchfield, BS, RHIA; April Post, BS^e

Abstract: **Background:** Hospital cancer registry data are used for a variety of aspects of patient care, yet one of the lesser used purposes of cancer registry data is to improve care coordination. **Objectives:** The purposes of this study were to assess hospital cancer registrars' perceptions of (1) the use of and quality of hospital cancer registry data for care coordination and other purposes; (2) the availability of all needed data for complete hospital cancer registry data collection; and (3) the data collection of COVID-19 effects on cancer patients. **Methods:** A survey was sent to hospital-based members of Cancer Registrars of Illinois between April and June 2020. Survey questions focused on current use and quality of hospital cancer registry data for care coordination as well as items related to COVID-19's effect on cancer patients. The focus of this study was hospital-based registrars, as they are the individuals collecting data directly from primary patient records. **Results:** While hospital cancer registry data are being used for many purposes including continuity of care, this study found that providers are not using hospital cancer registry data to its fullest extent. It was also found that hospital cancer registrars have collected valuable data on the impact that COVID-19 has had on cancer patients. **Conclusion:** Care coordination between providers is especially important for cancer patients who may see multiple providers and visit several facilities. This study found that the hospital cancer registry database contains extremely useful data for cancer patients and practitioners. Further, it was found that the hospital cancer registry is a source of valuable information regarding the impact that COVID-19 has had on cancer patients.

Key words: cancer registry data, care coordination, COVID-19 pandemic, data quality, data use, hospital cancer registry

Introduction

A preliminary review of the literature on cancer registry, care coordination, and cancer care coordination reveals a large number of studies;^{1,2} however, only a few directly discuss the comprehensive use of cancer registry data for care coordination. More importantly, the studies that discuss cancer registry data in care coordination are predominantly limited to the use of registry data in survivorship care plans.^{3,4} Other studies examine the use of electronic health record (EHR) data for care coordination,^{5,6} but these studies focus exclusively on EHR data and cancer registry data.⁷ No study specifically examines the quality and use of cancer registry data for care coordination purposes. This unexplored topic became particularly relevant in 2020, with the emphasis on care coordination for the complex and chronic diseases associated with cancer patients amid the effects of COVID-19.^{8,9} As the COVID-19 pandemic proceeded, it was noted that delays in cancer patient diagnosis, treatment, and follow-up were occurring. Patients were unable or afraid to seek health care for diagnosis, treatment, or follow-up. Health care facilities were canceling many elective procedures, including cancer screening and diagnostic tests. This upheaval in the health care system led to the need for documentation of the effect COVID-19 had on cancer patient care and care coordination.^{10,11,12}

Multiple, slightly different definitions of *care coordination* exist that reflect the importance of care among cancer patients.^{13,14} Care coordination is seen as a collaboration between patients, providers, and organizations that rely on the exchange and sharing of information. Care coordination has significant importance in terms of the provision of quality of care while attending to the reduction of duplicate tests, supporting appropriate and timely follow-up and support, and reviewing and capturing data. For the reasons mentioned above, the coordination of care is extremely important to cancer patients during active treatment, follow-up for chronic conditions, and in times such as the COVID-19 pandemic, when patients and providers face unique challenges.

Hospital cancer registry databases can be a valuable resource in cancer care coordination because they are designed to collect patient data beyond the acute phase of a disease. For example, data is collected at the time of the initial cancer diagnosis, during treatment, and beyond treatment through a systematic follow-up process. Cancer registry data include patient demographics, medical history, cancer characteristics, stages of disease, treatment, and outcomes.^{15,16} Such data can be very useful as the cancer patient moves beyond the acute phase of their disease into survivorship, follow-up, and recurrence. As various

^aIllinois State University, Normal, Illinois. ^bUniversity of Alabama at Birmingham, Birmingham, Alabama. ^cTemple University, Philadelphia, Pennsylvania.

^dSamford University, Birmingham, Alabama. ^eMcLean County Public Health Department, Bloomington, Illinois.

Address correspondence to Jennifer L Peterson, PhD, RHIA, CTR, Department of Health Sciences, Campus Box 5220, Illinois State University, Normal, IL 61709. (309) 438-7236. Email: jlpete2@ilstu.edu.

providers may be involved in such care, the hospital cancer registry database can serve as a repository and a resource for data needed by these various providers. Years after diagnosis, a patient with recurrent disease may see a new oncologist who can make treatment choices based on accurate current and past cancer data included in the registry, such as data from pathologists and cytologists. In addition, physicians and researchers can use such cancer data to learn more about the causes of cancer and to detect cancer earlier, thereby increasing the chances of finding a cure. Epidemiologists and researchers need cancer outcome data collected through patient follow-up to determine which treatments work and why some work and others do not. Since hospital cancer registries provide all these types of data, they are valuable research tools for those interested in the etiology, diagnosis, and treatment of cancer.¹⁷ Moreover, hospital registries are set up to perform monitoring and follow-up with patients, providers, and health care organizations, to create survivorship care plans, and to ensure accountability of high-quality data that can support cancer care coordination.

Cancer patients need strong care coordination for several reasons, including the complexity and nature of a cancer diagnosis; the sheer number of specialists and providers seen in the course of the care; the differences in geographic locations of patients, providers, and organizations; and the number and variety of health care settings.

Care coordination for cancer patients is needed and will require a wide variety of data beyond a single EHR, which does not contain all patient and disease data needed for collection in the cancer registry. Since data that is collected in the cancer registries is so important, and the integration of the EHR and registries has not been fully achieved, cancer registries should be the main source of data, and that data should be used to improve the quality of patient care. To meet the data needs of the cancer field, however, there must be a clear understanding of what data is collected by cancer registries and what data is needed.

The objective of this study is threefold: to assess hospital cancer registrar's perceptions of (1) the current use and quality of hospital cancer registry data for care coordination and other purposes; (2) the availability of all necessary data for complete hospital cancer registry data collection; and (3) the data collection of COVID-19's effects on cancer patients during the pandemic. Pertinent questions include:

- Are hospital cancer registries being used for care coordination?
- What are the uses for hospital cancer registry data?
- Do hospital cancer registrars have access to all data needed, and is it perceived to be complete and accurate?
- Are COVID-19 data being collected, and, if yes, what specific data is collected?

Methodology

Survey Development and Measurements

A survey was the main data collection approach used in this study. The survey design was based on general

information from various literature reviews. The publications reviewed included a variety of resources regarding cancer registry data use for cancer care,^{18,19,20,21,22} as well as patient care coordination measures.^{23,24}

Survey questions focused on the current use and quality of hospital cancer registry data for care coordination. Questions included assessment of hospital cancer registry data quality issues, especially related to data accessibility, completeness, and accuracy. Hospital cancer registry related data were categorized into 7 areas:

1. Patient information (demographics)
2. Disease information (stage, histology, morphology)
3. Comorbid conditions (COVID-19, other chronic conditions)
4. Genomic information (tests, results, consultation reports)
5. Laboratory values/vital signs/information (tumor markers, other pertinent labs, patient vitals)
6. Treatment information (agents, surgeries, radiation, start and stop dates, intent, termination reasons)
7. Outcome information (disease status, patient status, date of death or recurrence)

Based on the importance of the COVID-19 pandemic at the time of the survey, questions were included that addressed the collection of data related to the effects of COVID-19 on cancer patient care, cancer diagnosis, and treatment, including confirmed diagnosis of COVID-19 tests among cancer patients and information about cancer treatment delays due to the COVID-19 pandemic.

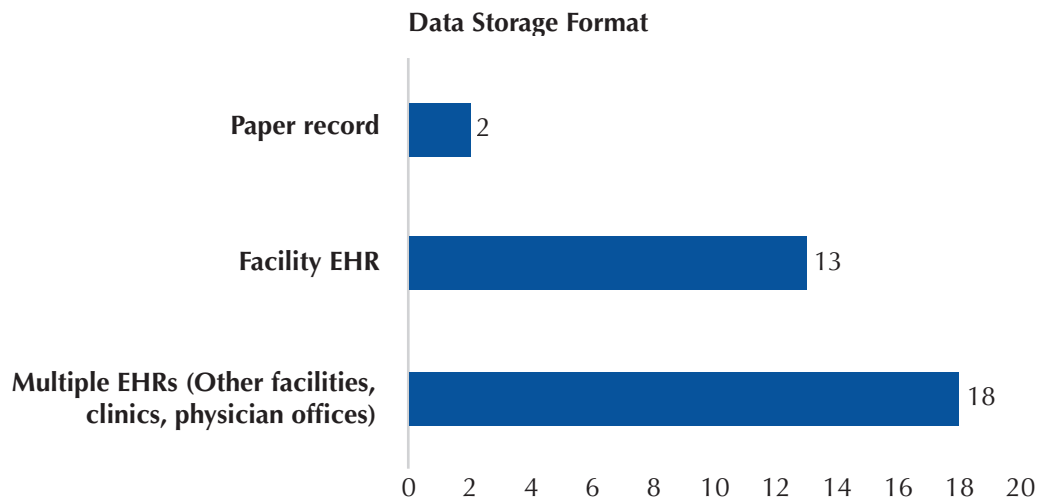
A self-designed survey was pilot tested among 5 local registrars to ensure the clarity of the questions and appropriate modifications were made. The study was reviewed and approved by the Illinois State University Institutional Review Board.

Study Participants and Survey Procedure

The membership list from the Cancer Registrars of Illinois was used for the survey distribution. The study participants included cancer registrars who were identified as hospital- or facility-based registrars from the state of Illinois. Members who worked for central registries and retired members were excluded, resulting in 89 individuals who were included as potential survey participants from the original list of 109 individuals. Hospital or facility-based registrars were selected as they are the individuals collecting data directly from the primary patient record. The focus of this study was on the hospital registrars' perceptions of the quality and use of hospital-based cancer registry data. Therefore, central registry personnel were not included as participants.

The survey was administered through Qualtrics and was distributed to the 89 initial potential participants on April 27, 2020. A total of 7 emails were undeliverable; thus, only 82 potential participants received the initial email invitation. Follow-up reminders were sent throughout month of May, and additional follow-up emails and phone calls were conducted in June 2020 to registrars who had not responded. Through this process, an additional 10 individuals were excluded from the initial email invitations,

Figure 1. Sources for Cancer Registry Data



which included 6 invalid emails and 4 potential participants who stated they were no longer in the cancer registry field. The final sample consists of 72 valid potential respondents with 28 survey respondents, resulting the final response rate of 38.9%. All survey participants consented to participate in the survey study.

Data Analysis

Data for this study was collected from April 2020 through June 2020. The data were analyzed using qualitative data analysis techniques and basic descriptive statistics. Data were initially analyzed through the Qualtrics report system. Quantitative data was analyzed using frequencies.

Qualitative data were analyzed using frequencies and modes while Likert scale questions were analyzed using frequencies of responses. Two of the authors analyzed the open-ended question responses using the constant comparative method. This was completed through initial manual coding of open-ended responses, followed by use of this coding to identify themes. General categories were first identified through constant comparative coding. Initially, broad categories were identified, followed by identification of specific details within the broad categories. Identification of these specific details led to the ability to discover themes. The 2 authors worked together to come to a consensus on the categories and themes. This allowed for the ability to organize the collected data into pieces that could be analyzed. After all the data were analyzed, categories and themes were integrated to provide an in-depth understanding of the interrelationships between the categories and themes. The coded data and themes were ultimately used to summarize the data and the overarching themes found. The constant comparative method enabled a thorough, organized approach to understanding the data.

Results

Respondent Characteristics

Of the 72 potential participants, 28 returned the survey; this yielded a response rate of 38.9%. One survey response was incomplete and was excluded from the analyses. The

majority (92.6%) of the respondents worked in an American College of Surgeons' Commission on Cancer (CoC)-accredited cancer program. More than half (51.9%) worked in an acute care hospital setting, one-third (33.3%) worked for a consulting service, and the remaining 14.8% worked in a cancer center. The majority of respondents who worked in a specific acute care hospital setting worked in facilities between 100 and 299 beds (40.0%) or 300 beds or more (50.0%). Only 10% worked in smaller facilities with fewer than 100 beds.

Most respondents (38.5%) held the job title of cancer registry abstractor; another 38.6% were titled cancer program coordinator. Other titles included cancer information specialist, consultant, data coordinator, director/manager/supervisor, lead cancer registrar, and cancer registrar. The term "registrar" is used in this analysis for simplicity. Because the surveys were distributed to the Cancer Registrars of Illinois members, the vast majority of respondents (74%) worked in a registry located in Illinois; respondents also worked in registries in Wisconsin, Colorado, California, Kentucky, and Massachusetts, or in multiple states.

Sources of Registry Information

The cancer registrar respondents indicated that they used multiple sources of information to collect hospital cancer registry data. Figure 1 shows the various sources that respondents used for obtaining the needed cancer registry data. Multiple electronic health records (EHRs) that were defined as EHRs with other facilities, clinics, and physician's offices were most often used to collect the complete information needed for a hospital cancer registry abstract. Almost half of respondents also used the facility EHR, with 2 reporting using paper records. An additional 2 respondents (not shown in the figure) reported using contact with physician offices and other hospitals.

When asked about the availability of specific cancer registry data items, the majority of registrars stated that most items were available through the paper record or the EHR. Some respondents, however, stated that some items

<i>Data item</i>	<i>Strongly agree % (n)</i>	<i>Agree % (n)</i>	<i>Neutral % (n)</i>	<i>Disagree % (n)</i>	<i>Strongly disagree % (n)</i>
Patient information	80.8 (21)	15.4 (4)	0.0 (0)	4.0 (1)	0
Disease information	50.0 (13)	34.6 (9)	7.7 (2)	7.7 (2)	0
Comorbid conditions	40.0 (10)	48.0 (12)	8.0 (2)	4.0 (1)	0
Genomic information	30.8 (8)	46.2 (12)	15.4 (4)	7.7 (2)	0
Laboratory values/vital signs	50.0 (13)	34.6 (9)	11.5 (3)	3.9 (1)	0
Treatment information	34.6 (9)	42.3 (11)	11.5 (3)	11.5 (3)	0
Outcome information	19.2 (5)	30.8 (8)	23.1 (6)	26.9 (7)	0

EHR, electronic health record. There were 26 responses except for comorbid conditions, which had 25 responses.

were not readily available. It is noted that outcome information was the data item for which the most respondents noted limited availability. Other data items that were noted to be available less often included treatment information and genomic information (Table 1).

Respondents offered many comments regarding the availability of the above data. These included many comments about the physician documentation, including the thoroughness of the documentation, as well as conflicting information between physicians. It was noted that patients may receive treatment at physician offices or other facilities, and obtaining information from these can be difficult. One registrar cited the need for “a strong relationship between the hospital and local cancer centers” as integral to obtaining complete information. Several registrars stated that obtaining all needed treatment information can be difficult and “very time consuming but critical.” The bottom line seemed to be best stated by the respondent who noted, “being able to locate data is inconsistent and not readily available.”

Respondents were slightly more positive in their perceptions about the data in the paper record and EHR being accurate. When asked if they perceived that specific cancer registry data items found in the original data source were accurate, most registrars stated that they felt that the items were accurate. The items that they felt to be less accurate somewhat mirrored the availability of the items; treatment information and genomic information were felt to be slightly less accurate. In addition, registrars noted that

disease information, such as stage, histology, and outcome information, were also less accurate (Table 2).

Most registrars noted that they perceived that most items were accurate, with one stating, “I can count on the documentation within the EHR to be accurate.” Another stated that they followed up with the appropriate department or physician to clarify any discrepancies. However, it was noted that some data items presented more challenges. The most commonly noted cancer registry data item that was perceived to be inaccurate was staging. It was noted that, “staging is not routinely documented in [the] EMR” and “physicians don’t take the time to fully stage a patient.” Documentation of disease status was noted to be inconsistent. Again, documentation regarding treatment was noted to be difficult to obtain; 1 respondent stated that “information on treatment provided at other facilities is [the] most challenging to capture.”

When asked if there were other data categories that were routinely collected in the hospital cancer registry, 3 of 21 (14.3%) respondents noted that clinical trial participation, weight/body mass index, and survivorship care plan receipt were included in their cancer registry data. When asked if additional data categories should be included in the cancer registry, 3 of 20 (15%) respondents stated that additional information should be included. Items that they felt should be included were further immune tests and DNA tests, data items specific to the National Accreditation Program for Breast Centers (NAPBC) and Cancer Program Practice Profile Reports (CP3R), and codes for why patients

<i>Data item</i>	<i>Strongly agree % (n)</i>	<i>Agree % (n)</i>	<i>Neutral % (n)</i>	<i>Disagree % (n)</i>	<i>Strongly disagree % (n)</i>
Patient information	40.0 (11)	52.0 (13)	0.0 (0)	4.0 (1)	0
Disease information	24.0 (6)	64.0 (16)	8.0 (2)	4.0 (1)	0
Comorbid conditions	32.0 (8)	60.0 (15)	0.0 (0)	8.0 (2)	0
Genomic information	36.0 (9)	52.0 (13)	4.0 (1)	8.0 (2)	0
Laboratory values/vital signs	44.0 (11)	48.0 (12)	4.0 (1)	4.0 (1)	0
Treatment information	32.0 (8)	56.0 (14)	4.0 (1)	8.0 (2)	0
Outcome information	24.0 (6)	52.0 (13)	12.0 (3)	12.0 (3)	0

There were 25 responses.

Figure 2: Reasons Data Requests Could Not be Filled

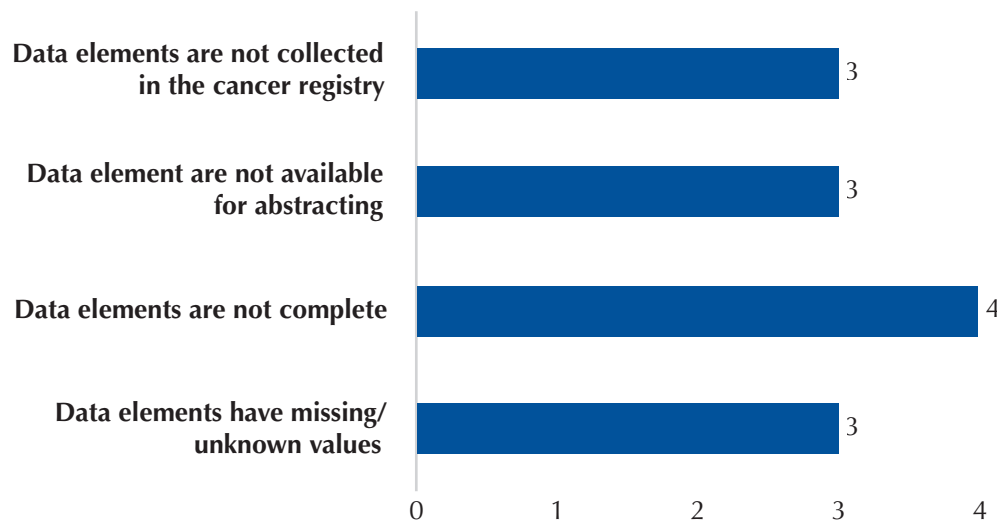


Table 3. Responses to “Your Cancer Registry Data are Used Regularly for the Following Purposes”

Use	Always % (n)	Often % (n)	Sometimes % (n)	Rarely % (n)	Never % (n)
Cancer conference patient ID	21.7 (5)	13.0 (3)	17.4 (4)	34.8 (8)	13.0 (3)
Cancer conference patient information	13.6 (3)	13.6 (3)	36.4 (8)	27.3 (6)	9.1 (2)
ID of patients for survivorship	34.8 (8)	39.1 (9)	4.4 (1)	4.4 (1)	17.4 (4)
Patient treatment summaries for survivorship	29.1 (6)	21.7 (5)	8.7 (2)	13.0 (3)	30.4 (7)
MD inquiries for treatment decisions	4.4 (1)	8.7 (2)	26.1 (6)	26.1 (6)	34.8 (8)
MD inquiries into patient status	8.7 (2)	4.4 (1)	26.1 (6)	26.1 (6)	34.8 (8)
Quality improvement studies	21.7 (5)	47.8 (11)	17.4 (4)	8.7 (2)	4.4 (1)
Benchmarking/ outcomes analysis	34.8 (8)	39.1 (9)	13.0 (3)	4.4 (1)	8.7 (2)
Physician or other research	8.7 (2)	34.8 (8)	30.4 (7)	17.4 (4)	8.7 (2)
Data analysis to inform MD practice	4.4 (1)	17.4 (4)	47.8 (11)	13.0 (3)	17.4 (4)
Admin reports for Staffing	8.7 (2)	30.4 (7)	21.7 (5)	13.0 (3)	26.1 (6)
Admin reports for resource allocation	13.0 (3)	30.4 (7)	21.7 (5)	17.4 (4)	17.4 (4)
Medical home seeking patient summary info	4.4 (1)	4.4 (1)	13.0 (3)	21.7 (5)	56.5 (13)
Facility-based community needs assessment	0.0 (0)	34.8 (8)	34.8 (8)	17.4 (4)	13.0 (3)
Community-based public health needs assessment	0.0 (0)	26.1 (6)	30.4 (7)	21.7 (5)	21.7 (5)
Continuity of patient care between providers	4.6 (1)	4.6 (1)	40.9 (9)	22.7 (5)	27.3 (6)
Continuity of care issues	8.7 (2)	34.8 (8)	47.8 (11)	4.4 (1)	4.4 (1)
ID of patients without complete treatment or ongoing follow-up care	0.0 (0)	56.5 (13)	30.4 (7)	4.35 (1)	8.7 (2)
Patient recall system beyond routine cancer registry follow-up	4.4 (1)	13.0 (3)	13.0 (3)	39.1 (9)	30.4 (7)

Admin, administration; ID, identification; info, information; MD, physician. Items addressed by 22 to 23 respondents; percentages vary based on total respondents.

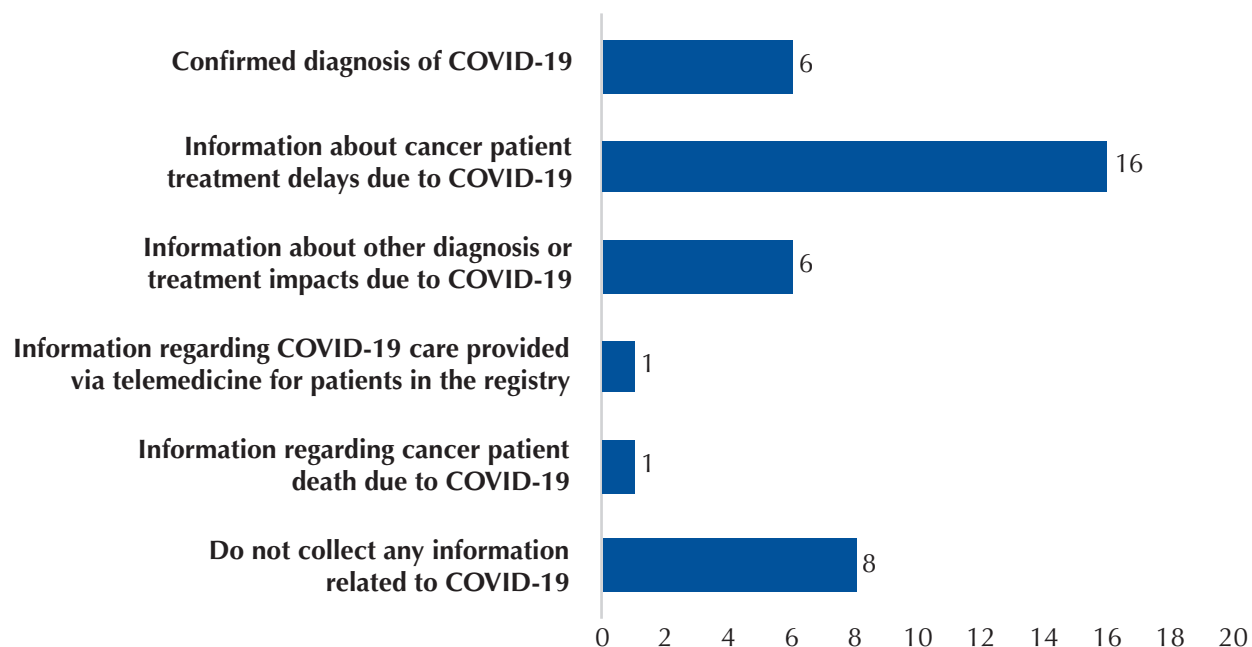
were not treated in the optimal time frame, as well as COVID-19 status and the impact this had on patients in terms of delayed diagnosis or treatment or change in treatment plan.

Use of Cancer Registry Data

Of the 26 registrars responding to this question, 22 (84.6%) stated that their cancer registry provides data for data requests; 2 (7.7%) stated that their registry did not

provide data and 2 (7.7%) were not sure. Of those who provided data for requests, 17 of 22 (77.3%) stated that they had all the data needed to fill the data requests. However, 5 (22.7%) stated that they did not. When asked why the data requests could not be filled, there were a variety of answers. Some of the reasons included that the data element is not available for abstracting, data elements are not collected or not completed in the cancer registry, and data elements have

Figure 3. COVID-19 Data Collected in Cancer Registries



missing or unknown values (Figure 2). Two respondents did not provide specific reasons.

Respondents were next asked about the use of hospital cancer registry data for specific purposes. The list of specific purposes ranged from common uses of cancer registry data, such as quality improvement studies, benchmarking and outcomes analysis, and identification of patients for survivorship treatment summaries, to more unusual uses such as physician or medical home inquiries for patient summary information. The use of such cancer registry data is summarized in Table 3.

Respondents were asked to provide specifics for the above noted items as appropriate. Respondents noted that hospital cancer registry data was used for some of the following purposes:

- Identification of patients for cancer conferences and to obtain appropriate diagnostic and treatment information
- Inquiries about patient status from physician staff
- Resident studies
- Medical group requests for informing physician practice
- Quality studies as required by various groups, including CoC, NAPBC, National Quality Forum, Quality Oncology Practice Initiative, American Society of Clinical Oncology, and the Qualified Clinical Data Registry
- Administrative purposes for staffing and resources
- Identification of community needs while working with outreach
- Identification of and follow-up of patients in collaboration with navigators for survivorship and ensuring complete treatment

Registry data was specifically noted to be rarely requested from medical homes seeking patient summary information.

When asked how frequently cancer conferences were held at their facilities, most respondents stated that they were held weekly (48.0%). Other responses were bimonthly (24.0%), monthly (4.0%), or unknown (24.0%). The unknown responses were primarily from consultants who worked at multiple facilities.

Cancer Registry COVID-19 Data

Due to the significant impact of COVID-19 on the health care system, respondents were asked to provide information regarding collection of COVID-19 data in the hospital registry, as well as their perception of the impact of COVID-19 on their cancer patients. The majority of respondents (63.0%) noted that they were collecting COVID-19 data in the cancer registry at the time of the survey. The most commonly collected data item was information about cancer patient treatment delays due to COVID-19, followed by confirmed diagnosis of COVID-19 and information about other diagnoses or treatments due to COVID-19. Eight respondents reported that they do not collect any information related to COVID-19 (Figure 3). In addition to the items listed in Figure 3, respondents were collecting data regarding cancer patient deaths due to COVID-19.

Respondents noted their perceptions of the many impacts and challenges related to the COVID-19 pandemic and cancer registry data collection. The most cited perceived impact was treatment delays for patients. Of the 21 responses to this question, 9 (42.9%) noted that they felt that COVID-19 treatment delays were having a major impact on patients. One registrar stated, "Patients are falling outside the optimal treatment window due to facility restrictions and patient decisions. Some have progression of disease due to delay. Some patients are afraid to come in for treatment." Other challenges noted were related to registrars working from home, which resulted in further challenges in obtaining all needed data. Respondents noted difficulty in

accessing all needed computer applications, the inability to obtain faxed reports, and the need for “a lot of coordination, equipment, and software meeting interfaces.”

Discussion

The data gathered from the cancer registry professional respondents enabled a deep understanding of the sources of hospital registry information, the availability and accuracy of that information, the use of the resultant cancer registry data, and the initial impacts of COVID-19 on hospital cancer registry data and cancer patients. The data that were provided were primarily from respondents working in American College of Surgeons' CoC facilities. The fact that most of the respondents were working in accredited programs adds to the credibility of the data findings.

In the information provided on the sources of registry information, it is noted that multiple EHRs are most commonly used and are needed to collect all the information required for a hospital cancer registry database. It is further noted, however, that there are several data items that are difficult to obtain and that not all information is available on EHRs to which registrars have access. Specifically, it was noted that treatment information is difficult to obtain and that physician documentation is not always complete in the EHR. Although it was noted that respondents felt that most types of cancer registry data were accurate within the EHR, there were some data items that registrars perceived to be less accurate, including treatment and genomic information as well as disease information and staging. Again, registrars pointed to the lack of information in the EHR or information being stored in EHRs to which they do not have access.

The availability and accuracy of cancer registry data is integral to the appropriate use of these data and decisions based on these data. The results of the survey point to the fact that registrars must use multiple EHRs and go outside the EHR systems to obtain thorough and accurate data for the registry. Respondents made it clear the need for accurate and complete data in the hospital registry. One respondent stated, “Our accreditation depends on high quality abstraction of required standard setter fields...to meet standards and report accurate outcomes.” It is clear that registrars are using all available resources to ensure the collection of complete and accurate data. However, it is also clear that complete cancer data on any given patient does not exist in any one EHR. This points to the value of the cancer registry database, in which all information regarding the cancer patient, their disease, diagnosis, treatment, and follow-up is available in one place.

While most respondents stated that their data was used for data requests, there were some respondents who could not provide the data requested based on the fact the data were not collected, the data were not available for abstracting, the data were not complete, or the data had missing or unknown values. Again, this points to the lack of complete information available to registrars for abstraction of cancer registry cases.

It was noted that hospital registry data in this study were used for many expected purposes such as survivorship, quality improvement, benchmarking or outcomes analysis,

physician or other research, to inform physician practice, administration reports for staffing and resource allocation, identification of patients without complete treatment or ongoing follow-up care, and for continuity of care issues such as time to diagnosis, time to treatment, and survivorship care. Hospital cancer registry data were used much more infrequently for purposes such as physician inquiries for treatment decisions or for patient status, medical homes seeking patient summary information, patient recall systems beyond routine cancer registry follow-up, and continuity of patient care between providers.

It is a positive sign that the respondents reported that cancer registry data are used for continuity of care in some areas. Some respondents reported working with navigators in their facilities in the areas of survivorship and ensuring complete treatment and ongoing care. This points to an understanding of the value of cancer registry data. However, the survey results show that many providers are not using the cancer registry database to their fullest extent. Based on the above findings indicating that cancer registry data include a complete picture of the cancer patient, their disease, treatment, and outcomes, it is clear that the hospital cancer registry database would be a valuable source of patient summary information for providers. In addition, it provides a rich source of information that could be used by providers to ensure continuity of care without duplication of services, while also serving to ensure appropriate follow-up. Cancer registry staff need to continue to work with cancer program staff and physicians to demonstrate the thorough and high-quality data available in the cancer registry and to encourage physicians to use these data for continuity of patient care purposes.

Due to the drastic impact that COVID-19 had on the health care system at the time of this survey, further information was collected regarding the effect that COVID-19 had on hospital cancer registry data collection as well as the registrars' perception of the effect COVID-19 had on cancer patient care. This survey found that most hospital registries were collecting COVID-19 data at the time of the survey. The data most frequently collected were in relation to treatment delays due to COVID-19. During the COVID-19 pandemic, many patients experienced delays in diagnosis and treatment due to stay-at-home orders, cancellation of all nonessential health care services, and public fear. This was also noted by the respondents as their perceived factor having the largest impact on cancer patients. Follow-up studies will be needed to see the long-term impact on disease progression, recurrence, and survival. Collection of COVID-19 data, including these delays, by hospital cancer registrars will be integral to the study of these long-term impacts. It is important that providers and administrators recognize that these data will be available in the hospital cancer registry when this issue is studied further.

One limitation of this study is that only members of Cancer Registrars of Illinois were included as participants. While this provides information regarding Illinois registrars' perceptions of the quality and use of hospital cancer registry information, this may not be representative of the perceptions of registrars across the United States. Another

limitation of this study was that the survey was completed during the height of the COVID-19 pandemic lockdown. Many individuals were working from home, including many registrars. Responses to the survey may have been influenced by this, either in regard to the participants who completed the survey or to their answers on the survey.

Conclusion

Care coordination between providers is integral to appropriate care for all patients. This coordination, however, becomes even more important for cancer patients who may be seeing multiple providers and using a number of difference facilities or institutions. Since cancer care is so often provided in multiple settings, no single EHR contains all the information the cancer registry does. While we are working to provide interoperability between systems, it is obvious that this is not yet completely successful. This study shows that the one place that holds all of the information about the cancer patient, as well as their diagnosis, treatment, and follow-up, is the cancer registry. While cancer registry data is used for a variety of purposes, it is still not being used to its fullest extent. Physicians and other providers could find a wealth of information in the registry that could be easily used for care coordination between providers. Cancer registrars should continue to showcase the data available in the registry to cancer program staff and physicians to encourage the use of the data for these purposes.

In the midst of the COVID-19 pandemic, it is clear that hospital cancer registries have quality documentation regarding the impact that COVID-19 has had on cancer patients. The majority of registrars in this survey have already started collecting COVID-19 data on their patients. This data will be invaluable in the future, when additional study will be needed to assess the long-term effects of COVID-19 on cancer patients.

This study has made it clear that registrars work extremely hard to ensure that they have the most thorough and accurate information in the cancer registry. These registries contain data from multiple sources in one place. These valuable data should be used to the fullest extent to provide patient care coordination and high-quality patient care.

References

- White M, Babcock F, Hayes N, et al. The history and use of cancer registry data by public health cancer control programs in the United States. *Cancer*. 2017;123:4969-4976. doi:10.1002/cncr.30905
- Gorin S, Haggstrom D, Han P, Fairfield K, Krebs P, Clauser S. Cancer care coordination: a systematic review and meta-analysis of over 30 years of empirical studies. *Ann Behav Med*. 2017;51(4):532-546. doi:10.1007/s12160-017-9876-2
- Salz T, Baxi S. Moving survivorship care plans forward: focus on care coordination. *Cancer Med*. 2016;5(7):1717-1722.
- Kvale E, Huang C, Meneses K, et al. Patient-centered support in the survivorship care transition: outcomes from the patient-owned survivorship care plan intervention. *Cancer*. 2016;122(20):3232-3242. doi:10.1002/cncr.30136
- Houser SH, Colquitt S, Clements K, Hart-Hester S. Impact of electronic health record usage on cancer registry systems in Alabama. *Perspect Health Inf Manag*. 2012;9(Spring):1f.
- Lee S, Jetelina K, Marks E, et al. Care coordination for complex cancer survivors in an integrated safety-net system: a study protocol. *BMC Cancer*. 2018;18(1):1204. doi:10.1186/s12885-018-5118-7
- Houser SH, Colquitt S, Clements K, Hart-Hester S. Impact of electronic health record usage on cancer registry systems in Alabama. *Perspect Health Inf Manag*. 2012;9(Spring):1f.
- Greene F. COVID-19 data and the cancer patient: a need for registry inclusion. *Ann Surg Oncol*. 2020;27(8):2571-2572. doi:10.1245/s10434-020-08554-y
- Raymond E, Thieblemont C, Alran S, Faivre S. Impact of the COVID-19 outbreak on the management of patients with cancer. *Target Oncol*. 2020;15(3):249-259. doi:10.1007/s11523-020-00721-1
- Morialiyage H, De Silva D, Ranasinghe W, et al. Cancer in lockdown: Impact of the COVID-19 pandemic on patients with cancer. *Oncologist*. 2021;26(2):e342-e344.
- Fujita K, Ito T, Saito Z, Kanai O, Nakatani K, Mio T. Impact of COVID-19 pandemic on lung cancer treatment scheduling. *Thoracic Cancer*. 2020;11:2983-2986.
- Al-Shamsi HE, Alhazzani W, Alhurairi A, et al. A practical approach to the management of cancer patients during the novel coronavirus disease 2019 (COVID-19) pandemic: an international collaborative group. *Oncologist*. 2020;25(6):e936-e945.
- McDonald KM, Sundaram V, Bravata DM, et al. Definitions of care coordination and related terms. In: *Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies (Vol. 7: Care Coordination)*. Agency for Healthcare Research and Quality (US); 2007. <https://www.ncbi.nlm.nih.gov/books/NBK44012/>
- McDonald KM, Schultz E, Albin L, et al. *Care Coordination Measures Atlas*. Agency for Healthcare Research and Quality (US); 2014:484. https://www.ahrq.gov/sites/default/files/publications/files/ccm_atlas.pdf
- Cancer registration & surveillance modules. National Cancer Institute SEER training modules website. https://training.seer.cancer.gov/modules_reg_surv.html
- Data Standards & Data Dictionary, Volume II. North American Association of Central Cancer Registries website. <https://www.naaccr.org/data-standards-data-dictionary/>
- Importance of cancer registry. National Cancer Institute SEER training modules website. <https://www.training.seer.cancer.gov/registration/registry/importance.html>
- Zachary I, Boren SA, et al. Information management in cancer registries: evaluating the needs for cancer data collection and cancer research. *Online J Public Health Inform*. 2015;7(2):e213. doi:10.5210/ojphi.v7i2.5664
- Mohammadzadeh Z, Ghazisaeedi M, et al. Systematic review of hospital based cancer registries (HBCRs): necessary tool to improve quality of care in cancer patients. *Asian Pac J Cancer Prev*. 2017;18(8):2027-2033. doi:10.22034/APJCP.2017.18.8.2027
- Campbell MK, Tessaro I, et al. 2011. Adult cancer survivorship care: experiences from the LIVESTRONG centers of excellence network. *J Cancer Surviv*. 2011;5(3):271-282.
- Hawley ST, Janz NK, Lillie SE, et al. Perceptions of care coordination in a population-based sample of diverse breast cancer patients. *Patient Educ Couns*. 2010;81(suppl):S34-S40.
- Look Hong NJ, Wright FC, et al. Multidisciplinary cancer conferences: exploring the attitudes of cancer care providers and administrators. *J Interprof Care*. 2009;23(6):599-610.
- Graetz I. *EHR Use and Care Coordination*. Agency for Healthcare Research and Quality (US); 2012.
- Agency for Healthcare Research and Quality (US). *Care Coordination Measures Database*. 2019.

Late-Stage Diagnosis and Cost of Colorectal Cancer Treatment in Two State Medicaid Programs

Sonja Hoover, MPP^a; Sujha Subramanian, PhD^a; Susan A. Sabatino, MD, MPH^b; Jaya S. Khushalani, PhD^b; Florence K. L. Tangka, PhD^b

Abstract: **Introduction:** To assess timing of Medicaid enrollment with late-stage colorectal cancer (CRC) diagnosis and estimate treatment costs by stage at diagnosis. **Methods:** We analyzed 2000–2009 California and Texas Medicaid data linked with cancer registry data. We assessed the association of Medicaid enrollment timing with late-stage colorectal cancer and estimated total and incremental 6-month treatment costs to Medicaid by stage using a noncancer comparison group matched on age group and sex. **Results:** Compared with Medicaid enrollment before diagnosis, enrolling after diagnosis was associated with late-stage diagnosis. Incremental per-person treatment costs were \$31,063, \$39,834, and \$47,161 for localized, regional, and distant stage in California, respectively; and \$28,701, \$38,212, and \$49,634 in Texas, respectively. **Discussion:** In California and Texas, Medicaid enrollment after CRC diagnosis was associated with later-stage disease and higher treatment costs. Facilitating timely and continuous Medicaid enrollment may lead to earlier stage at diagnosis, reduced costs, and improved outcomes.

Key words: cancer, cost, late-stage diagnosis, Medicaid, treatment

Introduction

Race and ethnicity have been documented as important factors in determining colorectal cancer (CRC) incidence and outcomes. Non-Hispanic Blacks or African Americans (hereafter referred to as Blacks or African Americans) compared with non-Hispanic Whites have a higher incidence of CRC, are diagnosed more often with distant-stage disease, and have a lower 5-year relative survival at any given stage.^{1–5} Hispanics experience disparities as well. They are diagnosed less often with early-stage CRC compared with non-Hispanic Whites, their incidence rates are increasing in comparison to non-Hispanic Whites, and they have worse outcomes when diagnosed with metastatic disease.^{2,6}

Medicaid is a vital source of health insurance for many low-income Americans, as it covers 1 in 7 adults aged 19–64 years.⁷ In addition, approximately 59% of Medicaid beneficiaries are Black or African American or another minority.⁸ Few studies have examined CRC treatment costs borne by Medicaid. Many studies on the cost of CRC treatment have derived estimates relevant to individuals 65 years or older covered by Medicare.^{9–12} Medicare costs, as they focus on treatments provided to older adults, may differ from Medicaid (or other insurance) costs for younger adults. Differences in CRC treatment have been reported between those insured by Medicaid versus Medicare.^{13,14} Because CRC treatment varies by stage at diagnosis, information on treatment cost stratified by stage for Medicaid patients

is important. Furthermore, given frequent discontinuity in Medicaid coverage among beneficiaries, it is important to understand the timing of Medicaid enrollment in relation to stage at diagnosis among enrollees.¹⁵ In this study, we analyze data for Medicaid beneficiaries aged ≤64 years from 2 states to examine the prevalence of late-stage CRC diagnosis in this population and determine whether stage at diagnosis is associated with timing of Medicaid enrollment or stratified by sociodemographic characteristics, as well as examining the costs of treating CRC by stage.

Methods

Data

We analyzed information about Medicaid beneficiaries living in California and Texas who were diagnosed with CRC during the years 2000–2009, between the ages of 21–64 years, and not dually eligible for Medicaid and Medicare. We selected the states of California and Texas because they had large cohorts of Medicaid beneficiaries and data were available. The institutional review board at RTI International, the California Health and Human Services Agency, and the Texas Department of State Health Services approved the research plan for this study.

To identify cases, California Cancer Registry (CCR) and Texas Cancer Registry (TCR) supplied identifiers in encrypted and password-protected files directly to the

^aRTI International, Waltham, Massachusetts. ^bCenters for Disease Control and Prevention, Atlanta, Georgia.

Address correspondence to Sonja Hoover, MPP, RTI International, 307 Waverley Oaks Road Suite 101, Waltham, MA 02452. Telephone: (781) 434-1722. Email: shoover@rti.org.

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

The collection of Texas cancer incidence data used in this study was supported by the Texas Department of State Health Services and Cancer Prevention Research Institute of Texas, as part of the statewide cancer reporting program and the CDC NPCR's Cooperative Agreement #1NU58DP006308. Funding support for Sujha Subramanian and Sonja Hoover was provided by the Centers for Disease Control and Prevention (CDC) (Contract No. 200-2008-27958 TO 36 to RTI International).

Centers for Medicare and Medicaid Services (CMS) for all those diagnosed with CRC from 2000 to 2009. CRC cases were identified with the following International Classification of Diseases for Oncology, second edition (ICD-O-2) codes: colon and rectum; colon excluding rectum, cecum (C180), appendix (C181), ascending colon (C182), hepatic flexure (C183), transverse colon (C184), splenic flexure (C185), descending colon (C186), sigmoid colon (C187), large intestine not otherwise specified (C188–C189, C260), rectum and rectosigmoid junction, rectosigmoid junction (C199), and rectum (C209). Due to delays in processing Medicaid claims data, this was the latest available information at the time of study initiation. CMS staff identified the cancer cohort for each state based on matches from the cancer registry data with the Medicaid Analytic eXtract (MAX) enrollment file. CMS sent RTI enrollment and claims files (personal summary, other therapy, long-term care, and prescription drug) for the matched patients and included a nonidentifiable patient identifier. RTI shared these files with CCR and TCR to obtain the relevant variables from the cancer registry database. CMS also sent data from 2000–2010 on beneficiaries for each state that did not have cancer to select an appropriate comparison group.

Our analytic sample consisted of individuals aged 21–64 years who were enrolled in Medicaid in either California or Texas. Beneficiaries who enrolled 3 months or more after their diagnosis of CRC were excluded from the study as we could not determine whether these individuals were enrolled in other plans prior to joining Medicaid. These beneficiaries may have had medical costs in the first 3 months of diagnosis that we could not capture because they were paid for by sources other than Medicaid. We excluded beneficiaries who were 65 years or older and those who were enrolled in both Medicaid and Medicare (dual enrollees) as we did not have complete utilization data to verify whether they were receiving Medicare services prior to official enrollment in Medicaid. For beneficiaries enrolled in both Medicare and Medicaid, Medicare is the primary payer, and we would not have been able to capture those costs.

Stage of Diagnosis at Enrollment Analysis

The Medicaid enrollment file contained beneficiary eligibility information, demographic characteristics, and monthly enrollment. Our sample consisted of beneficiaries who were enrolled prior to their CRC diagnosis who enrolled during the month of diagnosis and up to 2 months after diagnosis. Those who were enrolled prior to diagnosis were categorized as *enrolled prior to diagnosis* and those who enrolled within 2 months were included in the *enrolled after diagnosis* group. We extended the time frame to up to 2 months as many individuals attempt to enroll at the time of diagnosis; the length of time varies by state for the administrative processes of determining eligibility and finalizing Medicaid enrollment.¹⁶

Our overall sample consisted of 8,154 CRC patients in California and 4,044 CRC patients in Texas. We presented the estimates separately for California and Texas. We analyzed demographic (age, sex, race/ethnicity) and clinical

characteristics for patients enrolled in Medicaid before and after their cancer diagnosis. We reported race/ethnicity as non-Hispanic White, Hispanic (no specific race or multiple races), and Black or African American and other races/ethnicities combined. Black or African Americans made up about half of the latter group. Due to the small sample sizes of the groups of Black or African Americans and other races/ethnicities combined, we combined both groups. This allowed us to generate stable estimates, especially when examining specific time periods with smaller sample sizes. We ran logistic regressions to determine the probability of being diagnosed with late-stage disease. We defined *late-stage* as beneficiaries with cancer at regional or distant stage at diagnosis and compared with beneficiaries with cancer in in situ and localized stages as defined by Surveillance, Epidemiology, and End Results (SEER) Summary Stage.¹⁷ We compared beneficiaries enrolled after diagnosis vs those enrolled before diagnosis and controlled for age, sex, and the 3 broad race/ethnicity categories. For all analyses, *P* values < .05 were considered statistically significant.

Cancer and Noncancer Cohort Cost Assessment

To estimate accurate costs, the cancer cohort was limited to beneficiaries who were continuously enrolled in fee-for-service Medicaid for 6 months after diagnosis. Because this analysis focused on costs of CRC treatment, consistent with others, we also excluded beneficiaries who died within 6 months of diagnosis to avoid costs during end-of-life or terminal care.^{18,19} We created a noncancer matched cohort to compare costs with cancer patients. The noncancer cohort was similar to the cancer cohort and consisted of beneficiaries enrolled in Medicaid from 2000–2009 under the age of 65 years who were not dually eligible for Medicare. Each CRC patient was matched on age (aged 21–44 years, aged 45–64 years) and sex. Racial/ethnic group (non-Hispanic White, Hispanic, Black or African American and other races/ethnicities combined) was also included in the matching process when feasible; we were able to consistently use racial/ethnic group for matching the Texas cohorts. In addition, we also ensured that the follow-up period selected for the comparison case was the same as the cancer case to ensure that seasonal differences in cost did not impact our cost estimates. To accomplish this, we assigned a pseudo diagnosis date for comparison cases that was the same month and year as that of the diagnosis date of the cancer patient and also ensured that a continuous period of 6 months of fee-for-service enrollment from pseudo diagnosis date was available for cost estimation.

We included 2,850 CRC cases and 2,850 matched noncancer cases from California and 1,824 CRC cases and 1,824 noncancer cases from Texas to estimate cost of cancer treatment in the 6-month period after diagnosis. For both the cancer and noncancer cohorts, we calculated the total Medicaid costs and incremental costs of covered services from physician and outpatient visits, hospitalizations, prescription drugs, home health care, and long-term care facilities using the payment variable from each of the files. Incremental costs were calculated by subtracting the 6-month costs of noncancer patients (matched by age group

Table 1. Demographics and Clinical Characteristics of Medicaid Beneficiaries with Colorectal Cancer in California and Texas by Timing of Patient Enrollment in Medicaid

	California			Texas		
	Patients enrolled prior to diagnosis (n = 6,192)	Patients enrolled after diagnosis (n = 1,962)		Patients enrolled prior to diagnosis (n = 2,166)	Patients enrolled after diagnosis (n = 1,878)	
Age (mean y)	59.8	53.4	***	55.7	51.4	***
Sex (%)			***			***
Male	47.1	60.0		43.6	60.5	
Female	52.9	40.0		56.4	39.5	
Race/ethnicity (%)			***			*
Non-Hispanic White	33.7	37.0		35.6	35.2	
Hispanic	43.7	33.9		34.1	30.7	
Black or African American or other races/ethnicities combined	22.6	29.2		30.3	34.1	
Stage at diagnosis (%)			***			***
In situ	4.3	0.4		2.9	0.5	
Localized	31.4	10.8		29.0	11.4	
Regional	35.3	32.1		32.8	31.6	
Distant	24.7	54.5		22.8	50.6	
Unknown/unstaged	4.2	2.3		12.6	6.0	
Time from diagnosis to death (%)						
Died 0–6 months	13.8	19.4	***	15.8	15.2	
Died 7–12 months	6.8	11.6	***	8.7	12.3	***

*** $P < .001$; ** $P < .01$; * $P < .05$.

Statistical significance reported compares patients by timing of enrollment in Medicaid in each state.

and sex) with the total 6-month Medicaid costs for CRC patients. All costs are presented in 2018 dollars; cost of services for each year were inflated to 2018 estimates using the gross domestic product deflator.²⁰

Results

In Table 1, we compare demographic and clinical characteristics of patients enrolled prior to diagnosis and those enrolled after diagnosis. In California, all characteristics were statistically significantly associated with timing of enrollment. Patients enrolled prior to diagnosis were older than those enrolled after diagnosis (59.8 vs 53.4 years, respectively), and a higher percentage were female (52.9% vs 40.0%) and Hispanic (43.7% vs 33.9%). Patients enrolled prior to diagnosis were also diagnosed at an earlier stage: 35.7% of these patients were diagnosed at a localized stage or with in situ compared with 11.2% of patients enrolled after diagnosis. A higher percentage of patients enrolled after diagnosis, compared with patients enrolled prior to diagnosis, died within the first 6 months (19.4% vs 13.8%) and within the second 6 months after diagnosis (11.6% versus 6.8%).

Patients in Texas had similar characteristics. Patients enrolled before diagnosis compared with patients enrolled after diagnosis tended to be older (55.7 vs 51.4 years),

female (56.4% vs 39.5%), and Hispanic (34.1% vs 30.7%). They were also diagnosed at an earlier stage; however, there was also a higher percentage of patients enrolled prior to diagnosis who had an unknown stage or were unstaged (12.6% vs 6.0%). A higher percentage of patients enrolled after diagnosis died within the second 6 months (7–12 months) after diagnosis compared with patients enrolled prior to diagnosis (12.3% vs 8.7%).

We present the odds ratios of being diagnosed with late-stage CRC associated with year of diagnosis, enrollment before or after diagnosis, and demographics in Table 2. Age was a factor in late-stage diagnosis in California: overall from 2000–2009, Medicaid beneficiaries aged 40–49 years were 1.54 times higher odds of a late-stage CRC diagnosis compared with beneficiaries aged 60–64 years. California Medicaid beneficiaries aged 40–49 years showed similar trends in 2004–2006 and 2007–2009.

In California, patients diagnosed between 2000–2009 had 4.38-times higher odds of a late-stage CRC diagnosis if they enrolled after diagnosis than beforehand. The odds ratios increased from 3.67 for the group diagnosed in 2000–2003 to 5.50 for the group diagnosed in 2007–2009, although the confidence intervals overlapped, indicating that the increase may not be statistically significant. In Texas, patients diagnosed between 2000–2009 had 3.96-times

higher odds of being diagnosed late if they enrolled after diagnosis than before diagnosis. The odds ratios ranged from 3.78 for the group diagnosed in 2000–2003 to 4.52 for the group diagnosed in 2007–2009.

The total Medicaid per-person cost and incremental cancer treatment cost for the 6-month period after cancer diagnosis is shown by state and stage of diagnosis in Table 3. For both states, the total Medicaid cost increased by stage. In California, the total Medicaid per-person cost ranged from \$32,024 at the localized stage to \$47,832 for the distant stage. In Texas, the total Medicaid per-person cost ranged from \$31,414 in the localized stage to \$51,802 in the distant stage. The per-person incremental cancer treatment costs trended similarly as they increased with each stage. Per-person incremental costs ranged from \$31,063 for localized stage to \$47,161 for distant stage in California, and they ranged from \$28,701 for localized stage to \$49,634 for distant stage in Texas. Table 4 includes the descriptive characteristics of cancer patients and noncancer matches in California and Texas.

In Figure 1, we show the per-person incremental CRC treatment cost at 6 months by stage, type of service, and by state. In California, the per-person incremental treatment costs all increased as stage of diagnosis increased for ambulatory care services (\$9,137 for local, \$14,990 for regional, and \$20,117 for distant), hospital stays (\$17,498 for local, \$20,794 for regional, and \$23,052 for distant), and prescription drugs (ranged from \$2,466 for local, \$2,740 for regional, and \$3,112 for distant). Only long-term care did not. Long-term care services decreased as stage of diagnosis increased: \$1,962 per person at the local stage and \$880 per person at the distant stage.

In Texas, the per-person incremental costs of treatment in ambulatory care services (\$10,438 for local, \$17,132 for regional, and \$28,302 for distant stage) and hospital stays (\$15,839 for local, \$19,458 for regional and \$20,139 for distant stage) increased as stage of diagnosis increased. Incremental cost of treatment decreased as stage increased for long-term care services: \$1,550 in the local stage to \$568 in the distant stage. There was no pattern with prescription drugs: \$873 in the local stage, \$585 in the regional stage, and \$626 in the distant stage.

Discussion

Results from this study indicate that beneficiaries who were enrolled in Medicaid prior to diagnosis were more likely to be diagnosed at an earlier stage of CRC, whereas beneficiaries who enrolled after diagnosis were more likely to be diagnosed at a later stage. Beneficiaries may enroll in Medicaid after CRC diagnosis for a number of reasons. For example, a beneficiary might have been eligible for Medicaid but never enrolled in the program until after diagnosis.²¹ In addition, beneficiaries could have qualified for Medicaid under medically needy programs after they spent down income and depleted assets that did not allow them to qualify for the program at an earlier time.²² Our regressions showed that in both California and Texas, beneficiaries had nearly 4 times the odds of being diagnosed at a later stage of CRC if they were enrolled in Medicaid after diagnosis

compared with before diagnosis. These results indicate that continuous Medicaid enrollment is associated with earlier stage at diagnosis. Although our study provides specific evidence for CRC, prior studies have identified a similar pattern for breast cancer.^{23–25}

Medicaid beneficiaries are more likely to be diagnosed at a later stage compared with those with other types of insurance.^{26,27} One reason may be that many Medicaid beneficiaries experience a lack of continuity in Medicaid coverage,¹⁵ and this may be associated with a delayed diagnosis. Possible reasons for beneficiaries transitioning in and out of Medicaid (“churning”) include fluctuations in workplace insurance coverage or factors that may affect Medicaid eligibility, such as changes in income, residence, and family size, as well as administrative issues.^{15,28} One study analyzing the Medical Expenditure Panel Survey (MEPS) from 2000–2004 found that 2 million adults lose Medicaid each year; within 6 months of losing Medicaid, 17% have reenrolled, 34% had other coverage, but 49% remained without coverage.¹⁵ Additional studies were conducted to estimate potential changes in coverage and eligibility following passage of the Patient Protection and Affordable Care Act, and results indicated that churning (between Medicaid, health exchange plans, and no insurance) would continue.^{29,30} Because the Medicaid population generally has lower use of screening compared with the general population,^{31,32} this may be an additional reason why they are diagnosed with CRC at later stages. Further, results from one study of Medicaid beneficiaries in a managed care plan also indicated that Medicaid beneficiaries may not follow up with diagnostic colonoscopies after an abnormal screening.³³

In 2010, the cost of CRC care was the second highest by type of cancer, second only to breast cancer. At that time, the cost of CRC care was estimated at \$14.14 billion and was projected to increase to \$17.4 billion in 2020, again second to breast cancer.³⁴ Our analyses showed that total costs and incremental costs increased as beneficiaries were diagnosed at later stages, thus impacting the cost of CRC care. We note that, in each state, the total and incremental costs were nearly the same at each stage. This indicates that the comparison group of noncancer patients incurred very low Medicaid costs, possibly indicating that this group may have been healthier and had fewer medical costs.

Our analyses also indicated long-term care costs decreased by stage of diagnosis. Medicaid long-term care costs in this study were likely related to comorbid conditions and other underlying factors.³⁵ Beneficiaries with comorbid conditions may take longer to recover from cancer surgery and other treatments, and they incur higher costs likely due to complications.³⁶ Prescription costs were also generally lower as they were not a primary treatment option for CRC at the time of this study. Medicaid prescription costs in this study may reflect costs for treating chronic conditions as well as prescriptions to treat adverse effects and complications of cancer treatment.

There were limitations in the study. We defined beneficiaries “diagnosed at enrollment” as those beneficiaries who enrolled in Medicaid within 2 months of their month of

Table 2. Odds Ratios of Being Diagnosed with Late-Stage Colorectal Cancer for Medicaid Beneficiaries by Demographics and Timing of Enrollment																
	2000–2009				2000–2003				2004–2006				2007–2009			
	California (n = 7,847)		Texas (n = 3,658)		California (n = 2,354)		Texas (n = 1,236)		California (n = 2,444)		Texas (n = 1,124)		California (n = 3,049)		Texas (n = 1,298)	
Sex																
Male	1.00		1.00		1.00		1.00		1.00		1.00		1.00		1.00	
Female	0.98		1.02		0.86		0.86		1.11		1.18		0.99		1.04	
	(0.89–1.08)		(0.87–1.20)		(0.72–1.03)		(0.65–1.14)		(0.93–1.33)		(0.89–1.58)		(0.85–1.16)		(0.80–1.35)	
Age of diagnosis (y)																
40–49	1.54	***	1.22		1.29		0.83		1.56	***	1.36		1.73	***	1.47	
	(1.31–1.80)		(0.96–1.56)		(0.99–1.70)		(0.53–1.29)		(1.18–2.07)		(0.88–2.12)		(1.33–2.26)		(0.98–2.21)	
50–59	1.1		1.09		0.92		0.78		1.31	***	1.32		1.06		1.15	
	(0.99–1.23)		(0.90–1.33)		(0.75–1.13)		(0.53–1.15)		(1.08–1.60)		(0.93–1.86)		(0.89–1.26)		(0.85–1.55)	
60–64	1.00		1.00		1.00		1.00		1.00		1.00		1.00		1.00	
Race/ethnicity																
Non-Hispanic White	1.00		1.00		1.00		1.00		1.00		1.00		1.00		1.00	
Hispanic	0.9		0.83		0.71	***	0.74		0.93		0.88		1.06		0.91	
	(0.79–1.03)		(0.69–1.01)		(0.55–0.90)		(0.53–1.04)		(0.73–1.18)		(0.63–1.24)		(0.86–1.31)		(0.67–1.24)	
Black or African American or other races/ethnicities combined	0.91		1.05		0.91		1.00		0.84		1.02		0.97		1.16	
	(0.81–1.02)		(0.86–1.27)		(0.74–1.12)		(0.71–1.41)		(0.68–1.04)		(0.72–1.45)		(0.80–1.16)		(0.85–1.58)	
Enrolled in Medicaid at time of diagnosis	4.38	***	3.96	***	3.67	***	3.78	***	4.19	***	3.45	***	5.50	***	4.52	***
	(3.76–5.11)		(3.33–4.71)		(2.84–4.74)		(2.81–5.08)		(3.18–5.54)		(2.52–4.71)		(4.20–7.22)		(3.35–6.08)	

****P* < .0001
Reference groups: female, age of diagnosis 60–64 years, White race/ethnicity, not enrolled in Medicaid.
Dependent variable = late-stage diagnosis
Sample is Medicaid beneficiaries diagnosed 2000–2009.
Confidence intervals are contained in parentheses.

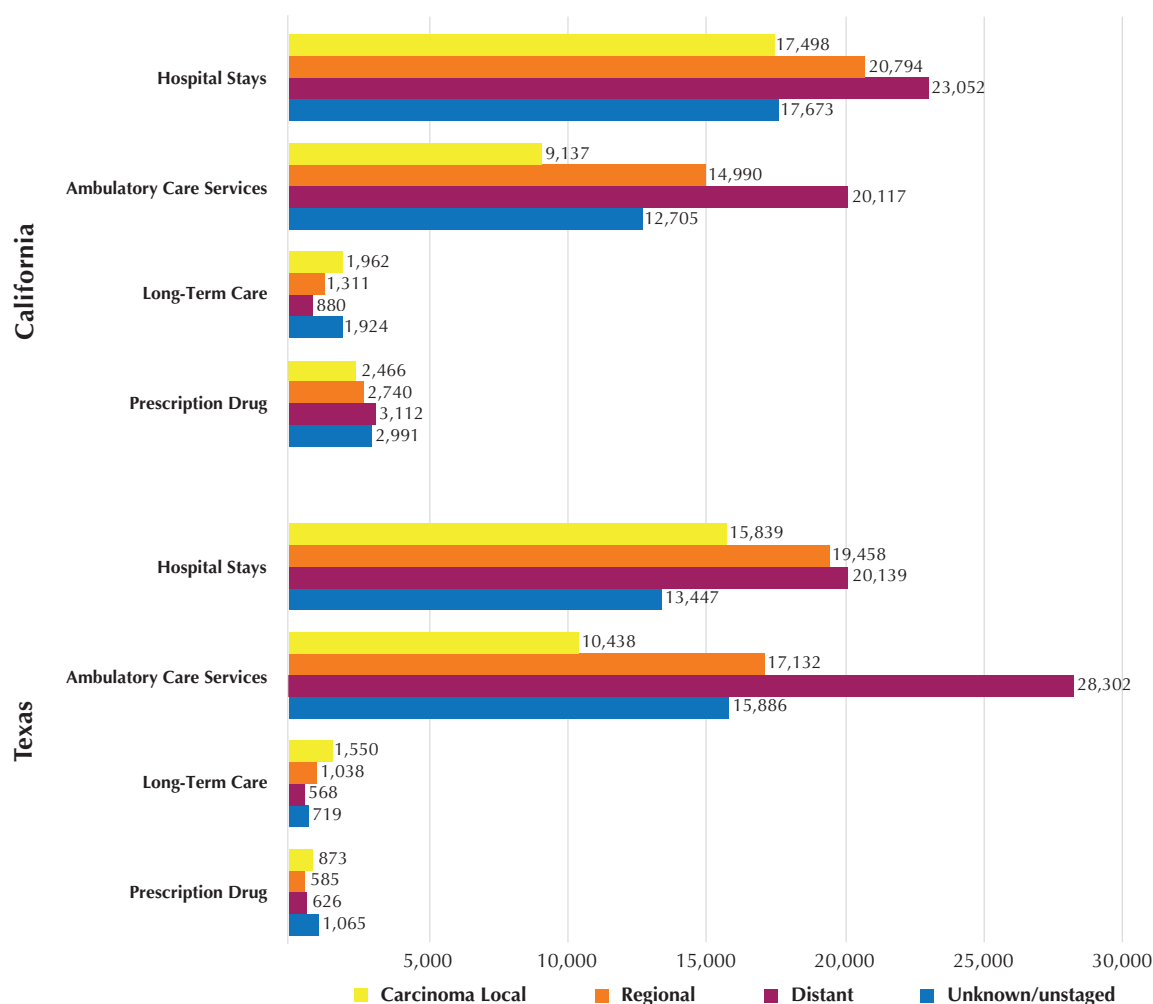
Table 3. Medicaid Total and Incremental 6-Month Per-Person Cost of Colorectal Cancer Treatment by State and Stage at Diagnosis in 2018 US Dollars				
	Total Medicaid cost ¹		Incremental cancer treatment cost ²	
	California	Texas	California	Texas
Stage at diagnosis				
Localized	32,024	31,414	31,063	28,701
	(29,207–34,841)	(28,905–33,922)	(28,208–33,919)	(26,157–31,244)
Regional	40,495	40,258	39,834	38,212
	(38,330–42,660)	(38,420–42,096)	(37,661–42,008)	(36,355–40,091)
Distant	47,832	51,802	47,161	49,634
	(45,534–50,131)	(49,459–54,144)	(44,866–49,457)	(47,247–52,021)
Unknown/unstaged	36,582	33,074	35,293	31,116
	(27,224–45,940)	(28,926–37,221)	(25,626–44,845)	(26,960–35,271)

¹Total Medicaid cost includes all costs for the 6-month period from cancer diagnosis.
²Incremental cancer treatment cost includes total Medicaid cost for colorectal cancer patients minus cost of noncancer patients matched by age group and sex. We included 2,942 and 1,858 colorectal cancer patients from California and Texas, respectively, who were matched with noncancer patients. In situ cases were excluded as there were very few cases and we were not able to report consistent or reliable estimates.

Table 4. Descriptive Characteristics of Cancer Patients and Noncancer Matches in California and Texas					
	California			Texas	
	Cancer patients (n = 2,942)	Noncancer matches (n = 2,942)		Cancer patients (n = 1,858)	Noncancer matches (n = 1,858)
Age, y (%)					
21–39	6.22	6.22		11.14	11.14
40–64	93.78	93.78		88.86	88.86
Sex (%)					
Male	52.21	52.21		50.81	50.81
Female	47.79	47.79		49.19	49.19
Race/ethnicity (%)			***		
Non-Hispanic White	39.73	34.06		38.32	38.32
Hispanic	37.19	35.76		27.83	27.83
Black or African American or races/ ethnicities combined	23.08	30.18		33.85	33.85

****P* < .001.

Figure 1. Per-Person Incremental Colorectal Cancer Treatment Cost (2018 US Dollars) at 6 Months by Stage and Type of Service



Note: In situ cases were excluded as there were very few cases, and we were not able to report consistent or reliable estimates.

diagnosis and, in doing so, we may have created a slightly less accurate period for analysis. Second, we excluded beneficiaries who did not remain enrolled 6 months after diagnosis, as we needed a continuously enrolled cohort to estimate cost. Treatments can take longer than 6 months, but many individuals unenroll from Medicaid after a limited period of enrollment.¹⁵ Although the time frame does not affect the 6-month cost estimates of cancer treatment reported in this study, it does impact the total cost to the Medicaid program. Only beneficiaries enrolled in fee-for-service Medicaid were included in the analyses, as we did not have complete information for those enrolled in Medicaid managed care. It is possible that there may be differences between beneficiaries enrolled in fee-for-service vs managed care. Additionally, we did not have a sufficient sample size for some racial/ethnic groups to support analyses by more specific race/ethnicity categories. The results may have limited generalizability to other state Medicaid programs as only 2 states were included in the analyses.

Although we used the same observation period to compute the cost of cancer and noncancer patients, we were unable to incorporate information on preexisting comorbidities, as many beneficiaries enrolled in Medicaid

after their diagnosis. Further, the data presented in this manuscript may not reflect current practice, and although costs are adjusted, they may be underestimated as costs may have increased at a higher rate than the adjustment. Lastly, because our focus was on treatment costs, we did not include end-of-life costs since these costs are significantly higher for cancer patients compared with noncancer patients,³⁷ and including these would have overestimated the net costs presented in this study.

Results from this study may have implications for the Medicaid program. As the study shows, treating CRC at later stages costs more than treating early stage disease. However, if CRC can be diagnosed early through CRC screening modalities, prognosis is better,³⁸ and, as suggested by study findings, costs are lower. The United States Preventive Services Taskforce recommends CRC screening for individuals aged 50–75 years.³⁸ Implementing evidence-based interventions to increase screening use, along with better understanding the reasons underlying differences in timing of enrollment, may help inform efforts to facilitate timely and continuous enrollment in Medicaid for those eligible and may lead to earlier stage at diagnosis, reduced costs, and improved outcomes for Medicaid beneficiaries.

Acknowledgements

We would like to acknowledge the California Health and Human Services Agency for the use of their data.

References

1. Hines RB, Markossian TW. Differences in late-stage diagnosis, treatment, and colorectal cancer-related death between rural and urban African Americans and whites in Georgia. *J Rural Health*. 2012;28(3):296-305. doi:10.1111/j.1748-0361.2011.00390.x
2. Jackson CS, Oman M, Patel AM, Vega KJ. Health disparities in colorectal cancer among racial and ethnic minorities in the United States. *J Gastrointest Oncol*. 2016;7(suppl 1):S32-S43. doi:10.3978/j.issn.2078-6891.2015.039
3. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin*. 2013;63(1):11-30. doi:10.3322/caac.21166
4. White A, Liu CC, Xia R, et al. Racial disparities and treatment trends in a large cohort of elderly African Americans and Caucasians with colorectal cancer, 1991 to 2002. *Cancer*. 2008;113(12):3400-3409. doi:10.1002/cncr.23924
5. Cronin KA, Lake AJ, Scott S, et al. Annual Report to the Nation on the Status of Cancer, part I: national cancer statistics. *Cancer*. 2018;124(13):2785-2800. doi:10.1002/cncr.31551
6. Koblinski J, Jandova J, Pandit V, Omesiete P, Nfonsam V. Disparities in colon and rectal cancer queried individually between Hispanics and Whites. *J Gastrointest Oncol*. 2019;10(4):632-640. doi:10.21037/jgo.2019.02.08
7. *Medicaid in the United States*. Henry J Kaiser Family Foundation; 2019. <http://files.kff.org/attachment/fact-sheet-medicaid-state-US>
8. Distribution of the nonelderly with Medicaid by race/ethnicity, 2017. Kaiser Family Foundation website. <https://www.kff.org/medicaid/state-indicator/distribution-by-raceethnicity-4/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>
9. Yabroff KR, Warren JL, Schrag D, et al. Comparison of approaches for estimating incidence costs of care for colorectal cancer patients. *Med Care*. 2009;47(7 suppl 1):S56-63. doi:10.1097/MLR.0b013e3181a4f482
10. Wright GE, Barlow WE, Green P, Baldwin LM, Taplin SH. Differences among the elderly in the treatment costs of colorectal cancer: how important is race? *Med Care*. 2007;45(5):420-430. doi:10.1097/01.mlr.0000257184.93944.80
11. Luo Z, Bradley CJ, Dahman BA, Gardiner JC. Colon cancer treatment costs for Medicare and dually eligible beneficiaries. *Health Care Financ Rev*. 2010;31(1):35-50.
12. Etzioni R, Ramsey SD, Berry K, Brown M. The impact of including future medical care costs when estimating the costs attributable to a disease: a colorectal cancer case study. *Health Econ*. 2001;10(3):245-256. doi:10.1002/hec.580
13. Sinclair AH, Schymura MJ, Boscoe FP, et al. Measuring colorectal cancer care quality for the publicly insured in New York State. *Cancer Med*. 2012;1(3):363-371. doi:10.1002/cam4.30
14. McGory ML, Zingmond DS, Sekeris E, Bastani R, Ko CY. A patient's race/ethnicity does not explain the underuse of appropriate adjuvant therapy in colorectal cancer. *Dis Colon Rectum*. 2006;49(3):319-329.
15. Sommers BD. Loss of health insurance among non-elderly adults in Medicaid. *J Gen Intern Med*. 2009;24(1):1-7. doi:10.1007/s11606-008-0792-9
16. *Medicaid MAGI and CHIP Application Processing Time Report*. Medicaid; 2019. <https://www.medicaid.gov/state-overviews/downloads/magi-and-chip-application-processing-time/magi-application-time-report-2019.pdf>
17. Young JL Jr, Roffers SD, Ries LAG, Fritz AG, Hurlbut AA, eds. *SEER Summary Staging Manual - 2000: Codes and Coding Instructions*. NIH Pub. No. 01-4969 ed. National Cancer Institute; 2001.
18. Subramanian S. Impact of Medicaid copayments on patients with cancer: lessons for Medicaid expansion under health reform. *Med Care*. 2011;49(9):842-847. doi:10.1097/MLR.0b013e31821b34db
19. Subramanian S, Trogdon J, Ekwueme DU, Gardner JG, Whitmire JT, Rao C. Cost of breast cancer treatment in Medicaid: implications for state programs providing coverage for low-income women. *Med Care*. 2011;49(1):89-95. doi:10.1097/MLR.0b013e3181f81c32
20. Data archive: national accounts (NIPA). US Department of Commerce Bureau of Economic Analysis website. <https://apps.bea.gov/histdata/histChildLevels.cfm?HMI=7>
21. Bradley CJ, Gardiner J, Given CW, Roberts C. Cancer, Medicaid enrollment, and survival disparities. *Cancer*. 2005;103(8):1712-1718. doi:10.1002/cncr.20954
22. Eligibility. Medicaid website. Accessed September 23, 2020. <https://www.medicaid.gov/medicaid/eligibility/index.html>
23. Subramanian S, Keating NL. Delays in breast cancer diagnosis after a state policy limiting Medicaid enrollment. *Cancer*. 2017;123(17):3219-3221. doi:10.1002/cncr.30769
24. Tarazi WW, Bradley CJ, Bear HD, Harless DW, Sabik LM. Impact of Medicaid disenrollment in Tennessee on breast cancer stage at diagnosis and treatment. *Cancer*. 2017;123(17):3312-3319. doi:10.1002/cncr.30771
25. Bradley CJ, Given CW, Roberts C. Correlates of late stage breast cancer and death in a Medicaid-insured population. *J Health Care Poor Underserved*. 2003;14(4):503-515. doi:10.1353/hpu.2010.0714
26. Andrew AS, Parker S, Anderson JC, et al. Risk factors for diagnosis of colorectal cancer at a late stage: a population-based study. *J Gen Intern Med*. 2018;33(12):2100-2105. doi:10.1007/s11606-018-4648-7
27. Sun W, Cheng M, Zhuang S, Qiu Z. Impact of insurance status on stage, treatment, and survival in patients with colorectal cancer: a population-based analysis. *Med Sci Monit*. 2019;25:2397-2418. doi:10.12659/msm.913282
28. Sommers BD, Gourevitch R, Maylone B, Blendon RJ, Epstein AM. Insurance churning rates for low-income adults under health reform: lower than expected but still harmful for many. *Health Aff (Millwood)*. 2016;35(10):1816-1824. doi:10.1377/hlthaff.2016.0455
29. Sommers BD, Rosenbaum S. Issues in health reform: how changes in eligibility may move millions back and forth between Medicaid and insurance exchanges. *Health Aff (Millwood)*. 2011;30(2):228-236. doi:10.1377/hlthaff.2010.1000
30. Sommers BD, Graves JA, Swartz K, Rosenbaum S. Medicaid and marketplace eligibility changes will occur often in all states; policy options can ease impact. *Health Aff (Millwood)*. 2014;33(4):700-707. doi:10.1377/hlthaff.2013.1023
31. de Moor JS, Cohen RA, Shapiro JA, et al. Colorectal cancer screening in the United States: trends from 2008 to 2015 and variation by health insurance coverage. *Prev Med*. 2018;112:199-206. doi:10.1016/j.ypmed.2018.05.001
32. Davis MM, Renfro S, Pham R, et al. Geographic and population-level disparities in colorectal cancer testing: a multilevel analysis of Medicaid and commercial claims data. *Prev Med*. 2017;101:44-52. doi:10.1016/j.ypmed.2017.05.001
33. Breen N, Skinner CS, Zheng Y, et al. Time to follow-up after colorectal cancer screening by health insurance type. *Am J Prev Med*. 2019;56(5):e143-e152. doi:10.1016/j.amepre.2019.01.005
34. Mariotto AB, Yabroff KR, Shao Y, Feuer EJ, Brown ML. Projections of the cost of cancer care in the United States: 2010-2020. *J Natl Cancer Inst*. 2011;103(2):117-128. doi:10.1093/jnci/djq495
35. Reaves EL, Musumeci MB; Kaiser Family Foundation. *Medicaid and Long-Term Services and Supports: A Primer*. The Kaiser Commission on Medicaid and the Uninsured; 2015. <https://files.kff.org/attachment/report-medicaid-and-long-term-services-and-supports-a-primer>
36. Subramanian S, Tangka FK, Sabatino SA, et al. Impact of chronic conditions on the cost of cancer care for Medicaid beneficiaries. *Medicare Medicaid Res Rev*. 2012;2(4):mmrr.002.04.a07. doi:10.5600/mmrr.002.04.a07
37. Tangka FK, Subramanian S, Sabatino SA, et al. End-of-life medical costs of Medicaid cancer patients. *Health Serv Res*. 2015;50(3):690-709. doi:10.1111/1475-6773.12259
38. United States Preventive Services Task Force; Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *JAMA*. 2016;315(23):2564-2575. doi:10.1001/jama.2016.5989

A Midwest Tri-State Study of Overall Survival in Ovarian Cancer with Adjuvant Chemotherapy

Sue-Min Lai, PhD, MS, MBA^a; Li Huang, MPH^a; Sarma Garimella, MBBS, MPH^a; John Keighley, PhD^b; Michele M. West, PhD^c

Abstract: **Background:** Overall survival associated with National Comprehensive Cancer Network (NCCN) adjuvant chemotherapy treatment guideline using population-based surveillance data is limited. This study examined overall survival and compliance to the NCCN guideline for adjuvant chemotherapy. **Methods:** The Midwest Ovarian Cancer Study was a collaborative project between 3 state cancer registries (Iowa, Kansas, and Missouri), Westat, and the Centers for Disease Control and Prevention. A standardized protocol was used to ascertain International Federation of Gynecology and Obstetrics (FIGO) stage-specific adjuvant chemotherapy. Primary epithelial ovarian cancers with FIGO stages IA/IB grade 3, IC, and II–IV with histologies 8000–8576 and 8930–9110 were included in this study. The Kaplan–Meier method was used to calculate survival functions. Adjusted hazard ratio (HR) was analyzed for all-cause mortality associated with NCCN compliance with adjuvant chemotherapy after adjusting for stage at diagnosis and comorbidity. **Results:** Sixty-nine percent (523 of 756 eligible) were compliant with NCCN guidelines. Compliance was significantly different by age at diagnosis and insurance type (both $P < .0001$). The overall survival was significantly different by age group, census tract median income, histologic subtype, and tumor grade (all $P < .0001$). The adjusted HR of noncompliance with adjuvant chemotherapy guideline was 3.2 (95% CI, 2.600–3.911). **Conclusions:** Better overall survival in patients who had received NCCN-recommended adjuvant chemotherapy was confirmed. **Impact:** The survival benefit was 7% higher over 4 years after diagnosis when considering FIGO stage-specific chemotherapy and the corresponding number of cycles. Using the chemotherapy data field that is collected by statewide cancer registries underestimated the overall survival.

Key words: National Comprehensive Cancer Network guideline compliance; National Program of Cancer Registries; ovarian cancer; overall survival, Surveillance, Epidemiology, and End Results Program

Introduction

Cancer of the ovary is ranked as the fifth-leading cause of cancer death among females, but the incidence of ovarian cancer is not among the top 5 leading cancers in women.¹ The high mortality rate resulting from cases being diagnosed at later stages can partially be attributed to not having an effective screening test or any unique symptoms associated with this cancer.^{2,4} Treatment for ovarian cancer starts with surgery to properly stage the cancer and to remove as much of the tumor as possible, followed by chemotherapy and targeted therapy for ovarian cancers of advanced stage. Survival is excellent when ovarian cancers are diagnosed at stage IA/IB and grade 1.^{5–10} Therefore, surgery alone is sufficient and recommended for tumors diagnosed at stage IA/IB and grade 1. Otherwise, the National Comprehensive Cancer Network (NCCN) recommends adjuvant chemotherapy for stages IA/IB grade 3,

IC, and II–IV.¹¹ The published literature regarding overall survival includes, for example, chemotherapy administered or levels of treatment compliance according to the NCCN clinical practice treatment guidelines and their effects on overall survival and recurrence.^{9–10,12–17} Findings on overall survival vary depending on population coverages (eg, Surveillance, Epidemiology, and End Results [SEER] Program, National Program of Cancer Registries [NPCR], clinical trials), tumor characteristics (eg, histologic subtype and staging methods) and definition of treatment used in the analysis.^{12–21} Population-based cancer registries collect whether chemotherapy was administered. However, detailed chemotherapy information—including International Federation of Gynecology and Obstetrics (FIGO) stage-specific chemotherapy information, such as the number of cycles that is necessary to address treatment compliance—are not collected.

^aKansas Cancer Registry, University of Kansas Medical Center, Kansas City, Kansas. ^bDepartment of Biostatistics, University of Kansas Medical Center, Kansas City, Kansas. ^cIowa Cancer Registry, Department of Epidemiology, The University of Iowa, Iowa City, Iowa.

Address correspondence to Sue-Min Lai, PhD, MS, MBA, University of Kansas Medical Center, Kansas City, KS 66160. Email: slai@kumc.edu. Telephone: (913) 588-2744.

This project was conducted by CDC, Westat, and the state cancer registries of Iowa, Kansas, and Missouri, and funded under CDC contract 200-2014-61258. The Iowa Cancer Registry is also funded in part with federal funds from NIH/NCI contract HHSN261201800012I and cancer center support grant NIH/NCI P30CA086862. The Kansas Cancer Registry is funded under CDC contract U58DP006273 and the Kansas Department of Health and Environment. The Missouri Cancer Registry is supported in part by a cooperative agreement between CDC and the Missouri Department of Health and Senior Services (DHSS) (U58DP006299-01/02) and a Surveillance Contract between DHSS and the University of Missouri.

Ovarian Cancer Treatment Study: Lisa L. Hunter, Charles F. Lynch, Michele M. West (Iowa Cancer Registry); Sue-Min Lai, Sarma Garimella, John Keighley, Li Huang (Kansas Cancer Registry); Jeannette Jackson-Thompson, Nancy Hunt Rold, Chester L. Schmaltz, Saba Yemane (Missouri Cancer Registry); Wilhelmina Ross, Diane Ng, Maricarmen Traverso-Ortiz (Westat); Jennifer M. Wike (CDC contractor); Trevor D. Thompson, Sun Hee Rim, Angela Moore, Sherri L. Stewart (CDC).

This study examines overall survival in ovarian cancer using data from the Midwest Ovarian Cancer Study,²² including patterns of adjuvant chemotherapy in women diagnosed with primary epithelial ovarian cancer with FIGO stages IA/IB grade 3, IC, and II–IV. Our study further evaluates overall survival associated with compliance of the NCCN guidelines for adjuvant chemotherapy recommendation from a large population cohort of patients. This study includes detailed information that is not part of standard data collection from the SEER Program and NPCR.

Methods

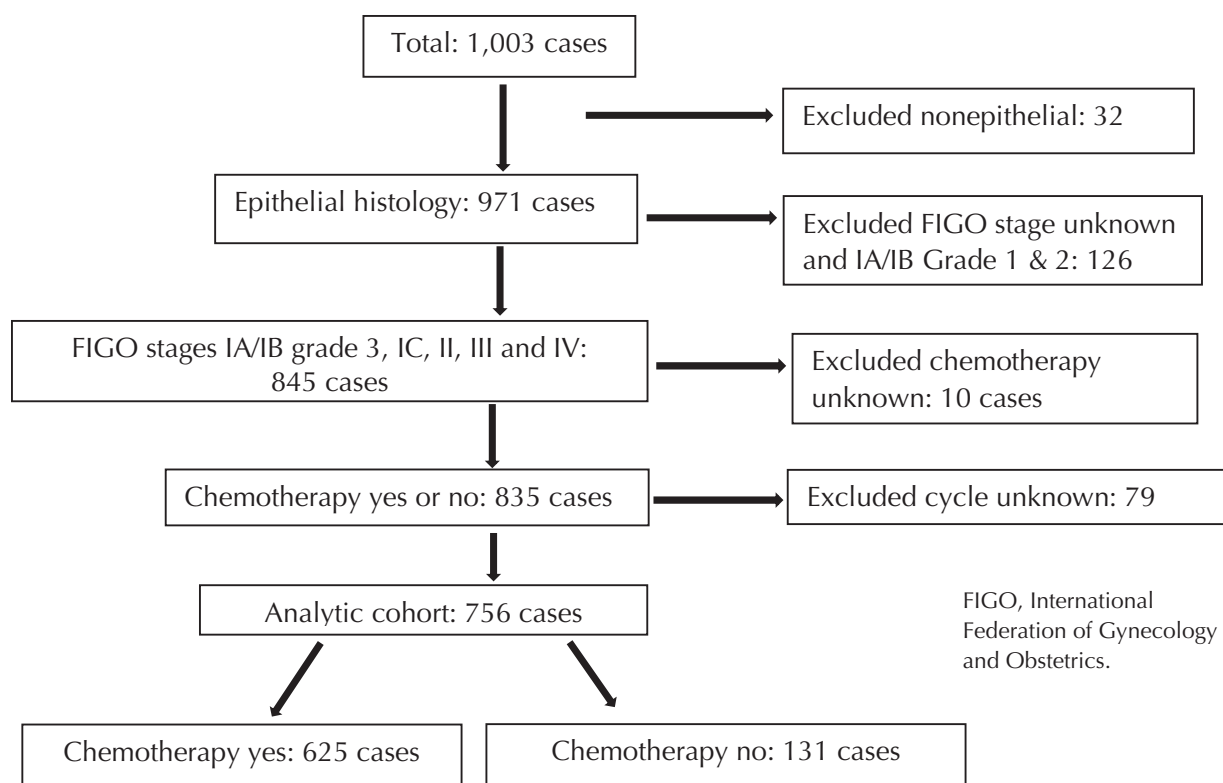
The Midwest Ovarian Cancer Study was a collaborative project between 3 Midwestern states (Iowa, Kansas, and Missouri), Westat, and the Centers for Disease Control and Prevention (CDC).²² Approvals from appropriate institutional review boards were obtained in 2017 for the 3 state cancer registries, Westat, and the CDC. The Midwest Ovarian Cancer Study included 1,003 primary invasive ovarian cancers that were diagnosed mostly from 2011 to 2012, with a few cases diagnosed in the later part of 2010, among women aged 18–89 years at the time of diagnosis. The primary sites codes were C56.9 (ovary), C57.0 (fallopian tube), and C48.1–C48.8 (peritoneum) with inclusion of histologies 8000–8576 and 8930–9110.²³ The inclusion criteria for this paper are primary epithelial ovarian cancers with FIGO stages IA/IB grade 3, IC, and II–IV (Figure 1).²⁴ The *International Classification of Diseases for Oncology*, third edition (ICD-O-3) codes were used to define histologic subtypes.²⁵

Use of primary adjuvant chemotherapy was evaluated according to the NCCN-recommended guidelines that were appropriate to cases diagnosed during years 2010–2012.¹¹ In addition to surgery, the recommended adjuvant chemotherapy for stages IA or IB grade 3 and IC was 3 to 6 cycles. A total of 6 to 8 cycles with intraperitoneal and/or intravenous administration were recommended for stages II–IV. Using these guidelines, patients were classified as *compliant* or *noncompliant* according to FIGO stage and grade. A data item to capture adjuvant chemotherapy routinely collected by most population-based cancer registries was also analyzed. The term *chemotherapy* is used throughout this manuscript to describe adjuvant chemotherapy drugs given to patients without documentation of type, cycles, and number of agents (single or multiagent chemotherapy administered).

Statistical Analyses

Independent variables included age, race, insurance/payer, census median income, census education, metro/nonmetro (or urbanicity) residence, and comorbidity score. *Census median income* is the median household income of the patient's census tract at the time of diagnosis. *Census education* is expressed as the percentage of residents that were high school graduates in the census tract where the patient lived at the time of diagnosis. A comorbidity score was calculated using the Charlson Comorbidity Index as adapted by the National Cancer Institute (NCI)²⁶ and presented on the SEER-Medicare website. The metro/nonmetro variable was obtained using the 2013 National

Figure 1. Midwest Ovarian Cancer Study Selection Criteria Flowchart



Center for Health Statistics urban-rural classification codes that were linked to cases based on state and county Federal Information Processing Standard (FIPS) codes. Vital status was ascertained by linkage of patients to state vital records and the National Death Index through December 31, 2018. The histologic subtypes were first individually tabulated, followed by a bivariate analysis after combining subtypes into 3 groupings due to small numbers in some subtypes.

Associations between categorical variables and compliance to the NCCN adjuvant chemotherapy guideline were examined using the χ^2 test and Fisher's exact test when appropriate. The primary end point of this study was all-cause mortality, which was calculated as the time interval between the day of diagnosis to the day of death or censored on December 31, 2016. The Kaplan-Meier survival curves of compliance vs noncompliance to the NCCN

Table 1. Demographic Characteristics and the National Comprehensive Cancer Network Adjuvant Chemotherapy Guideline Compliance: Midwest Tri-State Ovarian Cancer Study, 2017

<i>Characteristics</i>	<i>All cases n</i>	<i>Compliance n (%)</i>	<i>Noncompliance n (%)</i>	<i>P</i>
Total	756	523 (69.2)	233 (30.8)	
Median age (y)	65	62	74	<.0001
Age group (y)				<.0001
<50	86	69 (80.2)	17 (19.8)	
50–74	492	369 (75)	123 (25)	
≥75	178	85 (47.8)	93 (52.2)	
Race				.2934
White	714	497 (69.6)	217 (30.4)	
Others	42	26 (61.9)	16 (38.1)	
Insurance/Payer ¹				<.0001
Medicaid	38	26 (68.4)	12 (31.6)	
Medicare	381	228 (59.8)	153 (40.2)	
Military/Tricare/Veterans Affairs	11	5 (45.5)	6 (54.5)	
Not insured ² /self pay	46	36 (78.3)	10 (21.7)	
Private	273	225 (82.4)	48 (17.6)	
Urbanicity ¹				.352
Metro	503	354 (70.4)	149 (29.6)	
Nonmetro	252	169 (67.1)	83 (32.9)	
Census median income ¹				.2792
< \$51,000	364	245 (67.3)	119 (32.7)	
≥\$51,000	389	276 (71)	113 (29)	
Census high school graduates (%) ³				.5564
0–24	213	153 (71.8)	60 (28.2)	
25–49	528	359 (68)	169 (32)	
50–74	15	11 (73.3)	4 (26.7)	
75–100	0			
Comorbidity score				<.0001
0	573	423 (73.8)	150 (26.2)	
1	122	72 (59)	50 (41)	
2	32	20 (62.5)	12 (37.5)	
3	16	4 (25)	12 (75)	
≥4	13	4 (44.4)	9 (55.6)	

¹ Missing data in some cases.

² Eleven were not insured.

³ Percent of census tract residents that were high school graduates/equivalent) in where the patient lived at diagnosis.

Table 2. Tumor Characteristics and the National Comprehensive Cancer Network Adjuvant Chemotherapy Guidelines Compliance: Midwest Tri-State Ovarian Cancer Study, 2017

<i>Characteristics</i>	<i>All n</i>	<i>Compliance n (%)</i>	<i>Noncompliance n (%)</i>	<i>P</i>
Total	756	523 (69.2)	233 (30.8)	
Histologic tumor subtype				<.0001 ¹
Serous, mucinous, mixed, transitional cell carcinoma combined	560	424 (75.7)	136 (24.3)	
Serous	463	359 (77.5)	104 (22.5)	
Mucinous	30	19 (63.3)	11 (36.7)	
Mixed	64	45 (70.3)	19 (29.7)	
Transitional Cell	3	1 (33.3)	2 (66.7)	
Undifferentiated, other epithelial, and squamous ²	118	41 (34.7)	77 (65.3)	
Undifferentiated or other epithelial	117	40 (34.2)	77 (65.8)	
Squamous	1	1 (100)	0 (0)	
Clear cell and endometrioid carcinoma combined	78	58 (74.4)	20 (25.6)	
Clear cell	43	33 (76.7)	10 (23.3)	
Endometrioid	35	25 (71.4)	10 (28.6)	
International Federation of Gynecology and Obstetrics (FIGO) stage				.008
II, III, IV	643	433 (67.3)	210 (32.7)	
IA or IB grade 3	31	21 (67.7)	10 (32.3)	
IC	82	69 (84.1)	13 (15.9)	
Grade				<.0001
I and II	98	76 (77.6)	22 (22.4)	
III and IV	493	372 (75.5)	121 (24.5)	
Unknown	165	75 (45.5)	90 (54.5)	

¹ P-value was based on 3 combined histology subtypes.

² Included histologies 8010–8041, 8140, 8246–8260, 8410–8440, and 8574.

adjuvant chemotherapy guideline were analyzed using the log-rank test. The Cox proportional hazards regression was analyzed for all-cause mortality associated with NCCN compliance and other factors such as stage at diagnosis and the Charlson comorbidity score. The proportionality assumption was examined by creating interactions of the predictors and a function of survival time in the Cox model. The likelihood ratio test (LRT) was used to determine the level of significance. SAS version 9.4 for Windows was used for all analyses.

A sensitivity analysis was performed to evaluate if there was a bias in the estimated overall survival resulting from excluding 79 patients who had adjuvant chemotherapy treatment, but had unknown NCCN compliance status due to missing adjuvant chemotherapy cycles. Specifically, this sensitivity analysis calculated the cumulative survival functions for the following 3 groups of cases: 79 patients who were not included in the analytic dataset, 625 patients who had adjuvant chemotherapy and included in the analytic

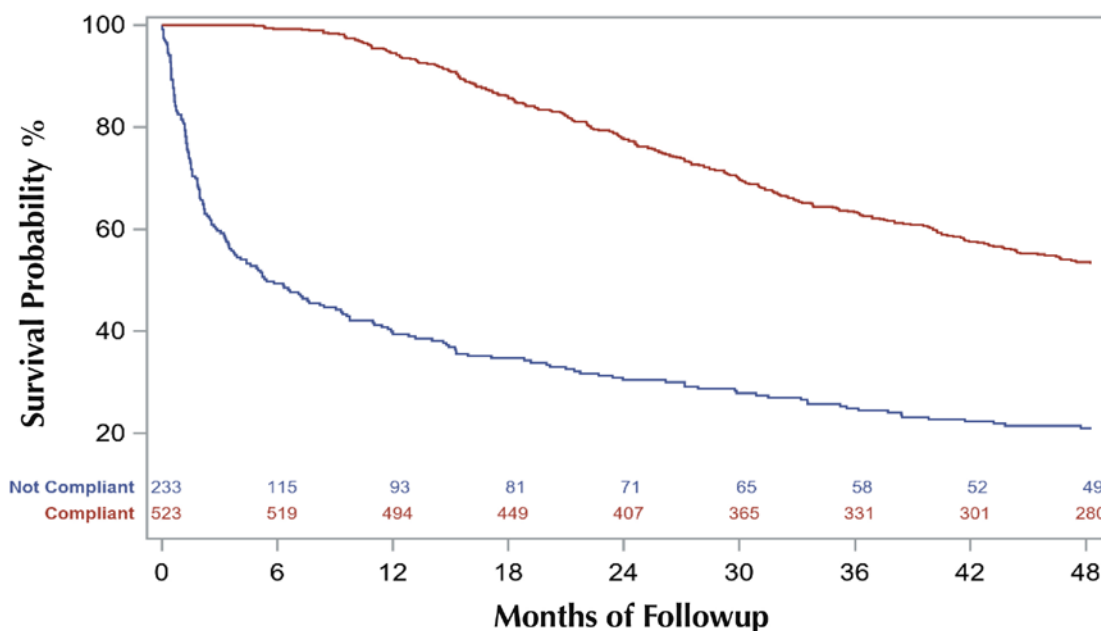
cohort, and 131 patients who had no adjuvant chemotherapy and were included in the analytic cohort of this study (Figure 1).

Results

A total of 756 patients who met the study criteria were included in the analysis (Figure 1). The demographic characteristics and NCCN guideline compliance to adjuvant chemotherapy treatment are shown in Table 1. Sixty-nine percent (523/756) of these women received the NCCN-recommended adjuvant chemotherapy treatment. The median age at the time of diagnosis for compliant patients and noncompliant patients was 62 years and 74 years, respectively ($P < .0001$). Age, insurance/payer status, and the comorbidity score at the time of ovarian cancer diagnosis were significantly different between cases who were compliant and those who were not compliant to the NCCN adjuvant chemotherapy guideline (all P values $< .0001$).

Ovarian cancer of serous carcinoma histology was the most common tumor type (61%; Table 2). The adjuvant

Figure 2. Overall Survival by National Comprehensive Cancer Network Adjuvant Chemotherapy Treatment Guideline Compliance



chemotherapy compliance status was significantly different across the 3 groupings of histologic type ($P < .0001$). The FIGO stage and tumor grade were also significantly different between those who were and who were not compliant with the adjuvant chemotherapy guideline ($P = .008$ and $P < .0002$, respectively).

The overall survival for the entire cohort of 756 women with ovarian cancer was 77.7%, 63.4%, 51.6%, and 46.7% at 12, 24, 36, and 48 months, respectively, after diagnosis. The median survival time was 37.9 months. The overall survival rates over 4 years were significantly better in compliance with NCCN guideline compared to noncompliance ($P < .0001$; Figure 2). The survival was 95%, 78%, 63%, and 53% at 12, 24, 36, and 48 months after diagnosis for those who were compliant with the NCCN adjuvant chemotherapy recommendation, while the survival was 40%, 31%, 25%, and 21% at 12, 24, 36, and 48 months after diagnosis for those who were not compliant with the NCCN guideline. The overall survival was also significantly different by histologic subtype, age group, and median income (all $P < .0001$). No difference was observed between those who lived in metro and nonmetro settings.

Table 3 shows the adjusted hazard ratio associated with all-cause mortality. All-cause mortality was significantly elevated in ovarian cancer patients noncompliant with the NCCN-recommended adjuvant chemotherapy (adjusted HR, 3.2; 95% CI, 2.60–3.91) after controlling for age group (LRT $P = .002$), race (LRT $P = .6151$), census median income (LRT $P = .0007$), comorbidity score (LRT $P = .0048$), histologic subtype (LRT $P < .0001$), FIGO stage (LRT $P < .0001$), and grade (LRT $P = .0013$).

The cumulative survival functions between the 625 ovarian cancer patients who were included in the analytic cohort and the 79 who were not included in the analytic cohort are shown in Figure 3 ($P < .0001$).

Discussion

The Midwest Ovarian Cancer Study is population-based with a coverage of ovarian cancers in 3 states and is in a unique position to bring additional insight beyond many of the previously published reports using the SEER^{7-8,19,28,29,33} and NPCR databases.^{4-6,32}

Both SEER and NPCR programs collect SEER Summary Stages 2000 (SS2000) at diagnosis, making it difficult to evaluate the patterns of treatment according to the NCCN-recommended treatment guidelines. The Midwest Ovarian Cancer Study group developed a standardized protocol and database to extract registry data items after validating with medical records, and also collected numerous additional information that is not routinely collected by statewide cancer registries (such as FIGO stage, reabstracted tumor grade, and cycles of adjuvant chemotherapies received by patients). Our study is able to examine the patterns of adjuvant chemotherapy treatment according to NCCN-recommended guidelines. We also evaluate overall survival from the perspective of the NCCN guideline but also from the perspective of statewide cancer registries that have a single data item to capture adjuvant chemotherapy. Our study showed a NCCN compliance rate of 69%. Women older than 75 years (47.8%) and with a higher comorbidity score were significantly less compliant to the NCCN-recommended adjuvant chemotherapy treatment compared with their counterparts. Women on Medicaid and Medicare were less compliant with the adjuvant chemotherapy guideline compared to women with private insurance or self-pay, though the observed lower compliance among Medicare beneficiaries was likely to be confounded with their age. Women who lived in metro/nonmetro, associated census tract level median incomes, and percent high-school graduates had similar levels of compliance to the recommended adjuvant chemotherapy. Effect of adjuvant chemotherapy

Table 3. Adjusted Hazard Ratios for All-Cause Mortality in Ovarian Cancer Patients (N = 756): Midwest Tri-State Ovarian Cancer Study, 2017

Factors	Adjusted Hazard Ratio ¹ (95% CI)	P
Age groups (y)		
<50	0.52 (0.342–0.777)	.0016
51–75	0.72 (0.580–0.900)	.0037
>75	Reference	
Race		
Non-White	1.13 (0.731–1.738)	.5875
White	Reference	
Census median income		
<\$51,000	1.38 (1.134–1.667)	.0012
≥\$51,000	Reference	
National Comprehensive Cancer Network guidelines		
Noncompliance	3.19 (2.600–3.911)	<.0001
Compliance	Reference	
Comorbidity score ²	1.16 (1.045–1.277)	.0047
Histologic subtype		
Clear cell, endometrioid	0.31 (0.183–0.514)	<.0001
Serous, mucinous, mixed, transitional cell	0.34 (0.256–0.447)	<.0001
Undifferentiated, other epithelial, squamous ³	Reference	
International Federation of Gynecology and Obstetrics (FIGO) stage		
IA/IB Grade 3	0.03 (0.004–0.215)	.0005
IC	0.14 (0.071–0.262)	<.0001
II, III, IV	Reference	
Grade		
I and II	Reference	
III and IV	1.71 (1.149–2.536)	.0081
Unknown	2.24 (1.448–3.476)	.0003

¹ Hazard ratio was adjusted for the factors included in the table.

² Test for trend.

³ Include histologies 8010-8041, 8140, 8246-8260, 8410-8440, and 8574.

treatment has been reported to be different with regard to race, ethnicity, payer status, and rural/urban (metro/nonmetro) in some studies,^{4,5,30,31,32} but were not in others.³³

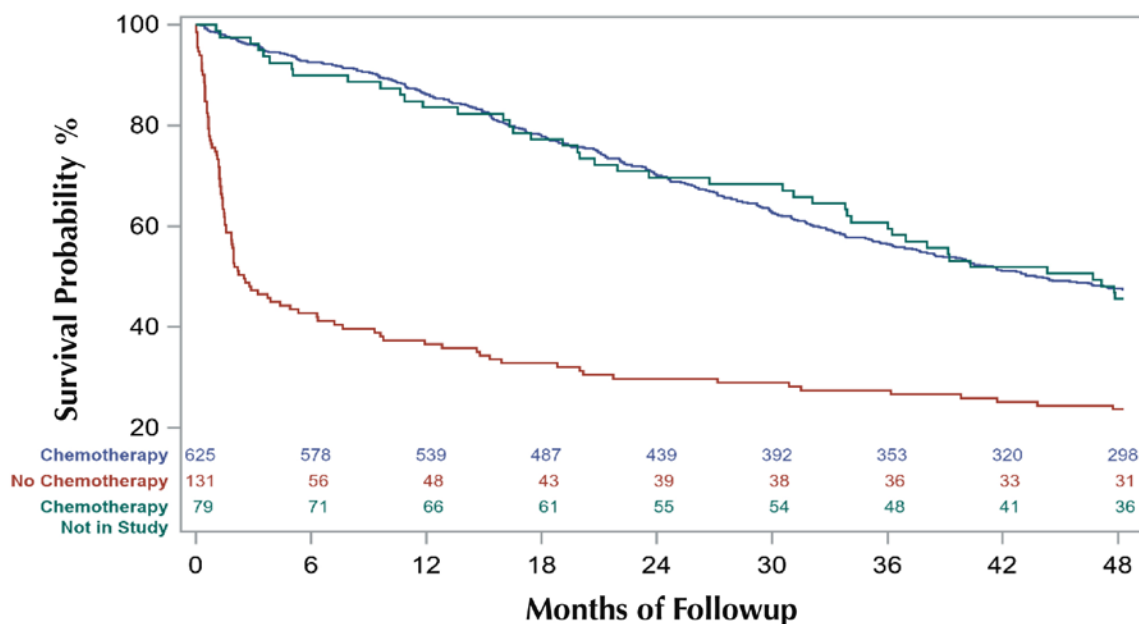
Overall survival in ovarian cancer patients have been reported. However, report of survival benefits from compliance to the NCCN treatment guidelines in ovarian cancer patients has been limited to coverages of hospital-based patients,^{9-10,12,13,15-17} and patients receiving care at an NCI cancer center.^{20,21,30} Our study is the only statewide cancer registries-based study that evaluates overall survival as stated previously. The survival was 95%, 78%, 63%, and 53% at 12, 24, 36, and 48 months after diagnosis for those who were compliant with the NCCN adjuvant chemotherapy recommendation, while the survival was 40%, 31%, 25%, and 21% at 12, 24, 36, and 48 months after diagnosis for those who were not compliant with the NCCN guideline. The adjusted hazard ratio of all causes of death is 3.19

times higher in NCCN treatment-noncompliant patients compared to NCCN-compliant patients after controlling for other factors such as age, comorbidities, FIGO stage, grade, and histologic subtypes.

With most US statewide cancer registries not collecting the necessary data items to evaluate compliance with the NCCN guideline, we reanalyzed the overall survival using the single data item *adjuvant chemotherapy* (Figure 3). Our analysis showed 7% higher survival rates over the 4 years following diagnosis when using NCCN adjuvant chemotherapy guideline compliance (Figure 2) relative to using the adjuvant chemotherapy data item (Figure 3). The additional 7% survival benefit over time is likely a result of more precise characterization of FIGO stage-specific adjuvant chemotherapy (NCCN) administered to the patients.

Our study cohort is representative of ovarian cancers at large with respect to demographics such as age, insurance/

Figure 3. Overall Survival by Receipt of Chemotherapy



payer, urbanicity, census median income, and medical comorbidities. Tumor characteristics are also consistent to the previous reports in regard to histologic subtype, FIGO stage, and grade. One limitation of our study is that the number of non-Whites was too small for a meaningful interpretation for these subgroups. Generalization of our survival finding to non-White cases in the United States may thus be limited. Lastly, the sensitivity analysis produced similar survival functions between the 2 subgroups of women ($n = 625$ and $n = 79$; Figure 3). The observed similar survival functions demonstrate no bias for the present study by not including 79 patients in the study cohort due to not having documented chemotherapy cycles received by patients.

In summary, FIGO-stage specific adjuvant chemotherapy greatly benefits overall survival in ovarian cancer patients who meet the NCCN adjuvant chemotherapy treatment guideline. Using the adjuvant chemotherapy data item that is available from all US statewide cancer registries is a good surrogate to the NCCN-recommended adjuvant chemotherapy treatment when evaluating the adjuvant chemotherapy benefit. The overall survival benefit from receiving the NCCN-recommended adjuvant chemotherapy for patients who meet the treatment guideline should at least be discussed at the time of treatment planning in spite of the adverse effects involved.

Acknowledgements

We would like to thank Sherri L. Stewart (CDC) for leading the publication committee, Jeff Steffens, BS, CTR, the MCR-ARC quality assurance staff, and the staff of facilities throughout the Missouri, Kansas, and Iowa central cancer registries for their dedication and desire for continuous quality improvement and submitting their reportable cases to central cancer registries.

References

1. Leading cancer cases and deaths, all races/ethnicities, male and female, 2017. Center for Disease Control and Prevention website. Published June 24, 2020. <https://gis.cdc.gov/cancer/USCS/DataViz.html>
2. Goff BA, Mandel LS, Drescher CW, et al. Development of an ovarian cancer symptom index: possibilities for earlier detection. *Cancer*. 2007;109(2):221-227. doi:10.1002/cncr.22371
3. Jacobs IJ, Menon U, Ryan A, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet*. 2016;387(10022):945-956. doi:10.1016/S0140-6736(15)01224-6
4. Stewart SL, Harewood R, Matz M, et al. Disparities in ovarian cancer survival in the United States (2001–2009): findings from the CONCORD-2 study. *Cancer*. 2017;123(suppl 24):5138-5159. doi:10.1002/cncr.31027
5. Cress RD, Chen YS, Morris CR, Petersen M, Leiserowitz GS. Characteristics of long-term survivors of epithelial ovarian cancer. *Obstet Gynecol*. 2015;126(3):491-497. doi:10.1097/AOG.0000000000000981
6. Cress RD, Bauer K, O'Malley CD, et al. Surgical staging of early stage epithelial ovarian cancer: results from the CDC-NPCR ovarian patterns of care study. *Gynecol Oncol*. 2011;121(1):94-99. doi:10.1016/j.ygyno.2010.12.359
7. Baldwin LA, Huang B, Miller RW, et al. Ten-year relative survival for epithelial ovarian cancer. *Obstet Gynecol*. 2012;120(3):612-618. doi:10.1097/AOG.0b013e318264f794
8. Du XL, Sun CC, Milam MR, Bodurka DC, Fang S. Ethnic differences in socioeconomic status, diagnosis, treatment, and survival among older women with epithelial ovarian cancer. *Int J Gynecol Cancer*. 2008;18(4):660-669. doi:10.1111/j.1525-1438.2007.01081.x
9. Cliby WA, Powell MA, Al-Hammadi N, et al. Ovarian cancer in the United States: contemporary patterns of care associated with improved survival. *Gynecol Oncol*. 2015;136(1):11-17. doi:10.1016/j.ygyno.2014.10.023
10. Nasioudis D, Latif NA, Simpkins F, et al. Adjuvant chemotherapy for early stage endometrioid ovarian carcinoma: an analysis of the National Cancer Data Base. *Gynecol Oncol*. 2020;156(2):315-319. doi:10.1016/j.ygyno.2019.11.125
11. Morgan RJ Jr, Alvarez RD, Armstrong DK, et al. Epithelial ovarian cancer. *J Natl Compr Canc Netw*. 2011;9(1):82-113. doi:10.6004/jnccn.2011.0008

12. Lee JY, Kim TH, Suh DH, et al. Impact of guideline adherence on patient outcomes in early-stage epithelial ovarian cancer. *Eur J Surg Oncol*. 2015;41(4):585-591. doi:10.1016/j.ejso.2015.01.006
13. Leiserowitz GS, Lin JF, Tergas AI, Cliby WA, Bristow RE. Factors predicting use of neoadjuvant chemotherapy compared with primary debulking surgery in advanced stage ovarian cancer: a National Cancer Database study. *Int J Gynecol Cancer*. 2017;27(4):675-83. doi:10.1097/IGC.0000000000000967
14. Kobal B, Noventa M, Cvjeticanin B, et al. Primary debulking surgery versus primary neoadjuvant chemotherapy for high grade advanced stage ovarian cancer: comparison of survivals. *Radiol Oncol*. 2018;52(3):307-319. doi:10.2478/raon-2018-0030
15. Seagle BL, Butler SK, Strohl AE, Nieves-Neira W, Shahabi S. Chemotherapy delay after primary debulking surgery for ovarian cancer. *Gynecol Oncol*. 2017;144(2):260-265. doi:10.1016/j.ygyno.2016.11.022
16. Nasioudis D, Haggerty AF, Giuntoli RL 2nd, et al. Adjuvant chemotherapy is not associated with a survival benefit for patients with early stage mucinous ovarian carcinoma. *Gynecol Oncol*. 2019;154(2):302-307. doi:10.1016/j.ygyno.2019.05.009
17. Rauh-Hain JA, Melamed A, Wright A, et al. Overall survival following neoadjuvant chemotherapy vs primary cytoreductive surgery in women with epithelial ovarian cancer: analysis of the National Cancer Database. *JAMA Oncol*. 2017;3(1):76-82. doi:10.1001/jamaoncol.2016.4411
18. Galvan-Turner VB, Chang J, Ziogas A, Bristow RE. Observed-to-expected ratio for adherence to treatment guidelines as a quality of care indicator for ovarian cancer. *Gynecol Oncol*. 2015;139(3):495-499. doi:10.1016/j.ygyno.2015.09.015
19. Warren JL, Harlan LC, Trimble EL, Stevens J, Grimes M, Cronin KA. Trends in the receipt of guideline care and survival for women with ovarian cancer: a population-based study. *Gynecol Oncol*. 2017;145(3):486-492. doi:10.1016/j.ygyno.2017.03.016
20. Petersen S, Shahiri P, Jewell A, et al. Disparities in ovarian cancer survival at the only NCI-designated cancer center in Kansas. *Am J Surg*. 2020;221(4):712-717. doi:10.1016/j.amjsurg.2020.12.009
21. Hildebrand JS, Wallace K, Graybill WS, Kelemen LE. Racial disparities in treatment and survival from ovarian cancer. *Cancer Epidemiol*. 2019;58:77-82. doi:10.1016/j.canep.2018.11.010
22. Weeks KS, Lynch CF, West M, et al. Impact of rurality on stage IV ovarian cancer at diagnosis: a Midwest cancer registry cohort study. *J Rural Health*. 2020;36(4):468-475. doi:10.1111/jrh.12419
23. Chen VW, Ruiz B, Killeen JL, Cote TR, Wu XC, Correa CN. Pathology and classification of ovarian tumors. *Cancer*. 2003;97(10 suppl):2631-2642. doi:10.1002/cncr.11345
24. Berek JS, Kehoe ST, Kumar L, Friedlander M. Cancer of the ovary, fallopian tube, and peritoneum. *Int J Gynaecol Obstet*. 2018;143(S2):59-78. doi:10.1002/ijgo.12614
25. Fritz A, Percy C, Jack A, et al. *International Classification of Diseases for Oncology*. 3rd ed. World Health Organization; 2000.
26. Mehta HB, Sura SD, Adhikari D, et al. Adapting the Elixhauser comorbidity index for cancer patients. *Cancer*. 2018;124(9):2018-2025. doi:10.1002/cncr.31269
27. Phippen NT, Barnett JC, Lowery WJ, Miller CR, Leath CA 3rd. Surgical outcomes and national comprehensive cancer network compliance in advanced ovarian cancer surgery in a low volume military treatment facility. *Gynecol Oncol*. 2013;131(1):158-162. doi:10.1016/j.ygyno.2013.07.001
28. Bristow RE, Chang J, Ziogas A, Anton Culver H. Adherence to treatment guidelines for ovarian cancer as a measure of quality of care. *Obstet Gynecol*. 2013;121(6):1226-1234.
29. Peres LC, Cushing-Haugen KL, Kobel M, et al. Invasive epithelial ovarian cancer survival by histotype and disease stage. *J Natl Cancer Inst*. 2019;111(1):60-68. doi:10.1093/jnci/djy071
30. Erickson BK, Martin JY, Shah MM, Straughn JM Jr, Leath CA III. Reasons for failure to deliver National Comprehensive Cancer Network (NCCN)-adherent care in the treatment of epithelial ovarian cancer at an NCCN cancer center. *Gynecol Oncol*. 2014;133(2):142-146.
31. Rauh-Hain JA, Melamed A, Schaps D, et al. Racial and ethnic disparities over time in the treatment and mortality of women with gynecological malignancies. *Gynecol Oncol*. 2018;149(1):4-11. doi:10.1016/j.ygyno.2017.12.006
32. Ore RM, Chen Q, DeSimone CP, et al. Population-based analysis of patient age and other disparities in the treatment of ovarian cancer in Central Appalachia and Kentucky. *South Med J*. 2018;111(6):333-341. doi:10.14423/SMJ.0000000000000821
33. Park J, Blackburn BE, Rowe K, et al. Rural-metropolitan disparities in ovarian cancer survival: a statewide population-based study. *Ann Epidemiol*. 2018;28(6):377-384. doi:10.1016/j.annepidem.2018.03.019

Attribute Associations as a Practical Approach to Account for Uncertainty in Residential Geolocation of Cancer Surveillance Data

Christian A. Klaus, MA^a; Carson Coggins, MS^b; Kevin A. Henry, PhD^{c,d}; Suzanne Bock, MPH^e; Dora Il'yasova, PhD^{e,f}

Abstract: This methodology article proposes a basic framework for assessing confidence in residential address through attribute sets of the tumor record that enable or modify spatiotemporal relationships in cancer surveillance data. A first step in assessing confidence for a statutory downstream data steward, like the Central Cancer Registry (CCR), is identifying sets of attributes whose domains are independently controlled by data stewards outside of the CCR. These include attribute sets that comprise the digital entities of *person*, *time*, and *place*. In this article, we describe the uncertainty in the geolocation of a cancer patient at the time of diagnosis, focusing on multiple stewardship of the cancer surveillance data. We also propose an approach to account for this uncertainty that is practical within the framework of existing cancer registry data coding, processing conventions, and legislative mandates for cancer surveillance.

Key words: attribute association, cancer, cancer cluster, geocoding, health informatics, population registry, residential geolocation, statutory stewardship

Introduction

As the term *cancer cluster* became part of the popular lexicon,¹ it has garnered attention toward uncovering the environmental origins of cancer clusters from the American public and legislators.² The public at large and the scientific community have a renewed interest in prospective spatial scanning for cancer clusters as a timely warning to the public of unexplained elevated incidence of cancer.³ However, due to operational and methodological constraints in interpreting spatial scan results, routine prospective cluster scanning has not been implemented.^{2,4,5} Our research addresses this concern by focusing on one of the most fundamental problems in the interpretation of scanned clusters—uncertainty that stems from geocoding when a proportion of the cases cannot be geocoded to an E911 address point (the address listed by the US emergency 911 system). Such uncertainty impacts confidence in incidence rates and their stability, not only in spatial cluster screening, but also in calculations of cancer incidence for relatively small enumeration areas. Specifically, we consider that cancer registries and epidemiologists working with cancer surveillance data are at best secondary stewards of the data comprising the 3 major epidemiologic entities of *person*, *time*, and *place*. In prior research, we proposed a template for quantified error probability of residential geolocation within the cancer surveillance data stream.⁶ In this article, we describe in more detail the uncertainty in the geolocation of a cancer patient at the time of diagnosis while focusing on multiple stewardship of the cancer surveillance data. We also propose an approach to capturing this uncertainty that is practical within the framework of existing data coding,

processing conventions, and legislative mandates for cancer surveillance.

A *cancer cluster*, by definition, is detected as “a greater-than-expected number of cancer cases that occurs within a group of people in a geographic area over a period of time.”⁷ By this definition, the detection of cancer clusters relates directly to the geographical specificity of cancer incidence rates. Historically, county-based cancer incidence rates have offered a tradeoff between geographic specificity and incidence estimate stability.^{8,9} Due to county-level statistical power, the county-based incidence estimates can provide comparability of rates across the larger geographical regions across states without a need to quantify the uncertainty in patient residence geolocation. Often, cancer clusters incur a greater geographical specificity, requiring higher-resolution scans.³ As the enumeration area size decreases (from county to subcounty, or even from large to small US states), so does the statistical power of detecting area-specific differences. Thus, at higher resolution, geocoding uncertainty is especially important in interpreting the results from routine prospective cluster screening.² In recent years, public pressure for higher-resolution scans has become evident,³ emphasizing the need for quantified confidence in residential geolocation.

In epidemiological surveillance, the monitoring of streams of patient abstracts includes evaluating attributes (or data fields) whose domains are controlled by different stewards. Currently, central cancer registries (CCRs) focus on screening for errors in surveillance data using tests of logical consistency called *edits*.^{10,11} Such screening detects errors in attributes sets whose domains and constraints are

^aNorth Carolina Central Cancer Registry, Raleigh, North Carolina. ^bPartner Software, Athens, Georgia. ^cDepartment of Geography and Urban Studies, Temple University, Philadelphia, Pennsylvania. ^dDivision of Cancer Prevention and Control, Fox Chase Cancer Center, Philadelphia, Pennsylvania. ^eDepartment of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia. ^fMTX Group Inc, Albany, New York.

The North Carolina Central Cancer Registry acknowledges the Centers for Disease Control and Prevention for its financial support under cooperative agreement NC U58 DP006281-04.

controlled by CCRs or their standard-setting organizations, with a few exceptions. However, these edits cannot detect errors in attributes controlled by non-CCR statutory and conditionally privileged stewards (ie, patients). Instead, for the data of other primary stewards, CCRs have only the option of assessing error probability. As curators of surveillance data, the disease registries bear the burden of error probability assessment. Thus, central to our investigation is the concept of stewardship in the cancer surveillance data stream.

Methods

Uncertainty Arising from Multiple Data Stewardship

We consider 2 primary levels of stewardship as relevant to the consolidated tumor records: *statutory* and *conditionally privileged (patients)*. Statutory stewards have been assigned centralized data stewardship through legislation, whereas patients are generally considered conditionally privileged stewards of their own demographic data. Statutory stewards are topic-specific and control certain data domains, as illustrated in Table 1. Statutory stewards also control the constraints of the attributes within the data domain (eg, US Social Security number and postal code cannot include more than 9 digits), so that confidence is enabled by domain constraints. In contrast, there are only limited domain constraints for patient demographic data, so confidence is enabled by record linkage verification alone. In addition, domain constraints for patient address are not always effective because patient address data may be more recent than geographic data that are available to registries for record linkage. In these cases, the registries must account for error probability without verification against all statutory address data.

Uncertainty arising from multiple data stewardship is best seen with attribute sets like patient identity and address at diagnosis. The US Social Security Administration controls the domain of the association of patient names, date of birth, and Social Security number. US address data stewards include city/county address data administrators, the US Postal Service (USPS), and the US Census. City/county address administrators steward house number, street name, pre- and post-direction (west, northeast, etc), and street type. USPS stewards the combination of these attributes plus postal code, whereas the Census stewards the combination of all of these attributes plus enumeration area (except county). At any given time, for the vast majority of addresses, these stewards concur on allowable combinations of address components. However, when data are consolidated over time, the address attributes from different stewards can disagree. One consequence of multiple stewardship in the cancer surveillance data stream is that quantified confidence in record linkage can decrease, resulting in inconsistent data quality.

Based on our analysis of the 2020 consolidated tumor record,¹² multiple stewardship is seen in approximately 9% of attributes, as these records are authored at least in part by patients, or their medical facility proxies, with minimal or no constraints. Multiple stewardship is best illustrated

in residential addresses where patients can produce the most inconsistency in the address domain. To correct this inconsistency, data from conditionally privileged stewards (patients) must be periodically reconciled with the official version(s) of those data from statutory stewards. However, in this reconciliation process, the burden falls on the patient to correct or fill in the gaps in their clinical data through online portals to their clinical information; otherwise, it falls on disease registries to correct or fill gaps through record linkage. Compounding this burden, the United States does not have unified regulations of ownership or stewardship of personal information,¹³ thereby removing incentives for patients to monitor repositories of their demographic data for quality. This situation further strengthens the need to quantify confidence in the error probability of cancer incidence data.

Uncertainty Arising from Changes to Address or Patient Demographic Data Over Time

Uncertainty can also arise by the way the CCRs obtain patient and address data. Rather than receiving continuous streams, CCRs receive and release data on patients and addresses from upstream stewards in discrete vintages (time snapshots). As a result, there is a time lag between the most recent vintage of non-CCR stewarded data and incorporating this information into CCRs' released data. The gap between the release of the CCR data and its use by researchers guarantees that at least some cancer cases will be impacted. At the time of data consumption, there is no guarantee that upstream stewards have not changed the address and patient identity domains so that attributes that were once concordant become discordant. In other words, CCRs are not necessarily a party to the future consensus of permissible domains of data from upstream stewards. Although the percentage of cases impacted by such discordance is low, the total number of cases enumerates in hundreds of thousands, if not millions. Thus, this discordance creates a need to offset expectations of attribute set certainty against future contradicting evidence.

Attribute Associations and Evaluation of Error Probability in Residential Geolocation

Statistical measures of association evaluate the strength of association between attributes for which no relationship is a priori assumed. In contrast, in the context of cancer surveillance record linkage, we are interested in the presumed associations between the attributes. Prior research on attribute and tuple uncertainty has not focused on error probability constrained by conditionally privileged or statutory stewards. Database uncertainty is modeled at the levels of tuple (attribute sets) and attributes,¹⁴ with a clear divergence between models assuming tuple independence¹⁵ versus dependencies within the data.¹⁶ Our analysis assumes attribute set (or tuple) independence between rows in the consolidated tumor record and, as a result, focuses on attributes only.

In the context of unknown prior data stewardship, attribute uncertainty is a core concern within database schemas for modeling relations between and within entities and their

Table 1. Attribute Sets Designated for Quantification of Error Probability via Evaluation of Attribute Associations (AAs)					
Attribute set no.	Description	Primary steward	Key components	AA error probability	Selected standards for CCRs
1	Tumor	Imaging/biopsy laboratory or physician (>98% of cases); patient (<2% of cases)	Primary site, histology, laterality	Probability that the tumor is not identified as distinct from other tumors	Primary site [Item #400] ¹² Histologic type ICDO3 [Item #577] ¹² Laterality [Item #410] ¹²
2	Date of diagnosis	Imaging/biopsy laboratory or physician (>98% of cases); patient (<2% of cases)	Month, day, year	Probability that the date of diagnosis cannot be specified to a day	Date of diagnosis [Item #390] ¹²
3	Patient identity	US address stewards, patient demographic data stewards,* and person-centric demographic data stewards	Government-issued identification number, name(s), date of birth	Probability that patient has not been positively identified	Social Security number [Item #2320] ¹² Date of birth [Item #240] ¹²
4	Address at diagnosis	US address stewards and person-centric demographic data stewards*	Patient address, date of diagnosis	Probability that chosen address at diagnosis does not span date of diagnosis	Addr_at_Dx_No_Street [Item #2330] ¹² Addr_at_Dx_City [Item #70] ¹² Addr_at_Dx_Postal_Code [Item #100] ¹²
5	E911 address point or E911 address point proxy	E911 administrative staff	Address, postal code, postal locality, county, state, country, geocode	Probability that 1 and only 1 E911 address point is not associated with patient address	Section 3.4.2 Placement of address point based on a parcel ²⁴
6	Place of residence building unit	E911 administrative staff	Address, postal code, postal locality, county, state, country, geocode and sub-address	Probability that patient sub-address specific address is not associated with 1 and only 1 E911 address point specific to digital building footprint or building image	Section 3.5 Address point placement for sub-addresses ²⁴
7	Place of residence building	Public safety data administrative staff	Address, postal code, postal locality, county, state, country, geocode and digital building footprint or building image	Probability that patient address is not associated with 1 and only 1 digital building footprint or building image	Section 3.4.4.2 Manual placement ²⁴
8	Place of residence parcel	Property tax parcel data administrative staff	Address, postal code, postal locality, county, state, country, discrete area, identification number(s)	Probability that parcel associated with patient address is not associated with patient address geocode	Section 3.4.2 Placement of address point based on a parcel ²⁴
9	Nested enumeration area 1 (US Census block group)	US Census	Polygonal topology, identification number, discrete area	Probability that the E911 address point record linkage candidates of equivalent likelihood† for a given patient address are associated with more than 1 Enumeration Area 1	Census Block Group 2000 [Item #362] ¹² Census Block Group 2010 [Item #363] ¹²

Table 1, cont. Attribute Sets Designated for Quantification of Error Probability via Evaluation of Attribute Associations (AAs)

<i>Attribute set no.</i>	<i>Description</i>	<i>Primary steward</i>	<i>Key components</i>	<i>AA error probability</i>	<i>Selected standards for CCRs</i>
10	Nested enumeration area 2 (US Census tract)	US Census	Polygonal topology, identification number, discrete area	Probability that the E911 address point record linkage candidates of equivalent likelihood for a given patient address are associated with more than 1 Enumeration Area 2	Census Tract 2000 [Item #362] ¹² Census Tract 2010 [Item #363] ¹²
11	Nested enumeration area 3 (US county)	County geospatial data administrative staff	Polygonal topology, identification number, discrete area	Probability that the E911 address point record linkage candidates of equivalent likelihood for a given patient address are associated with more than 1 Enumeration Area 3	County at DX Geocode 2000 [Item #95] ¹² County at DX Geocode 2010 [Item #96] ¹²
12	Nested enumeration area 4 (US state)	State geospatial data administrative staff	Polygonal topology, identification number, discrete area	Probability that the E911 address point record linkage candidates of equivalent likelihood for a given patient address are associated with more than 1 Enumeration Area 4	State at DX Geocode 2000 [Item #82] ¹² State at DX Geocode 2010 [Item #93] ¹²
13	Nested enumeration area 5 (Country)	Federal geospatial data administrative staff	Polygonal topology, identification number, discrete area	Probability that the E911 address point record linkage candidates of equivalent likelihood for a given patient address are associated with more than 1 Enumeration Area 5	Addr at DX Country [Item #102] ¹²
14	Incorporated area(s)	Municipal planning staff	Polygonal topology, identification number, discrete area	Probability that the E911 address point record linkage candidates of equivalent likelihood for a given patient address are associated with more than 1 Municipal Area	Addr at DX City [Item #70] ¹²
15	Derived area: postal locality	Third party vendors of US Postal Service data	Polygonal topology, identification number, discrete area	Probability that the E911 address point record linkage candidates of equivalent likelihood for a given patient address are associated with more than 1 vendor maintained postal code area	Addr at DX Postal Code [Item #100] ¹²

CCR, certified cancer registrar.

*Vendors of personal demographic and behavior data that CCRs are allowed access to in the United States under the Fair Credit Reporting and Gramm-Leach Bliley Acts.

[†]E911 address point record linkage candidates of equivalent likelihood are geocoding record linkage candidates that might be matched to patient address with equivalent justification, based on patient or reference data address missing address components, or with incorrect values for address components.

attributes.¹⁷ This is especially true for modeling the propagation of attribute errors such as inaccuracy, incompleteness, and lack of timely corrections.¹⁸ An attribute association (AA) connotes the potential for uncertainty in relationships between attributes due to multiple stewardship in a data stream. AA, as the term suggests, is a set of attributes (within a tuple) with potentially any relationship between them, including concordance, discordance, or uncertainty based on cross-field domain constraints. These relationships arise from the conferral of association by statutory or conditionally privileged stewardship. Often, AAs present hierarchies in which adding or subtracting an attribute results in incrementally more specific or general sets. Thus, an AA effectively connects metadata to well-established data fields, providing an opportunity to quantify their certainty. As such, an AA presents a default functional unit to account for error probability in residential geolocation while balancing the competing needs for the granularity of analysis and controlling the expense and replicability of metadata capture across CCRs.

AA-specific metadata for patient identity and address are especially useful for managing expectations of confidence in attribute sets. These attribute sets are presented in Table 1 according to the hierarchical structure of the cancer surveillance data. The direct verification of all attributes is extremely time- and labor-consuming, because as the scale of cases increases, verifying attributes can become unfeasible based on the resources of the CCR. Thus, there has been an effort among researchers to optimize geocoding systems.¹⁹ As case counts rise above a certain threshold, AA-specific metadata maintain a record of data integrity or lack thereof. Therefore, confidence in cancer surveillance also becomes a function of confidence in AA. Previously, we mapped the epidemiologic concepts of *person*, *time*, and *place* to 5 core attribute sets that enable or modify spatiotemporal relationships in cancer surveillance data.⁶ Here, we add geospatial granularity to our list by expanding it to 15 core sets (Table 1). We deem these 15 AAs represent a more complete evaluation of error probability of residential geolocation in the consolidated tumor record. Developing this revised list of attribute sets, we pursued the following principles:

1. To capture attribute sets that allow record linkage to other data sets and thereby their own verification and subsequent analytic capability.
2. To encompass all levels of granularity of spatial analysis commonly employed in epidemiologic investigation. The availability of data with different spatial specificity levels has been limited, while they appear crucial for prospective scanning of cancer clusters.
3. To balance the competing needs of analysis granularity while controlling the expense and replicability of metadata capture across CCRs.
4. To minimize the number of attribute sets while adhering to principles 1 through 3 to assist in forging an enduring consensus about the CCRs' practices in evaluating the error probability for residential geolocation.

There are 3 levels of granularity in the designated attribute sets: entity (eg, tumor characteristics, date at diagnosis), attribute (E911 address), and attribute component

(eg, address components such as postal code, geocode). Of the 3, we focus on the entity level because it is best for forging long-term consensus. As such, a future change in the attribute sets within an entity will be easily incorporated in the proposed framework as a modification while not requiring a more fundamental change to a shared understanding of the entity itself.

The designated 15 AAs (Table 1) arise from several data streams: medical records (attribute sets 1–4), residence (attribute sets 5–8), census (attribute sets 9–13), municipal area(s) (attribute set 14), and postal locality (attribute set 15). The first 2 AAs define tumor characteristics and the date of diagnosis, which can serve to account for the error probability in the identification of a distinct tumor and a specific date of diagnosis. The third AA accounts for patient identity, which is central to the linkage between the diagnosis, date of diagnosis, and address at diagnosis (ie, fourth AA). The error probability in patient identity and address at diagnosis is accounted through linking these attributes to similar attributes in other data sources.

Finally, we define AAs for different levels of residential geolocation spanning from E911 address point (or its proxy) to the parcel, building, and building units. For each level, we describe the data source and error probability. A general outline of how we estimated error probabilities for E911 address point and enumeration areas is found in Klaus et al, 2015, and the error probabilities for other AA will be described in future articles.⁶ We also note that our inclusion of E911 address points, parcels, and digital building footprints is made fiscally practical by the availability in North Carolina of statewide E911 address points and property tax parcels^{20,21}—with standardized attributes for each—as well as statewide building footprints.²²

In addition to residential geolocation, we describe 5 AAs that are important in evaluating error probability when residential address is nested within an administrative geographical division, such as US census block and tract, county, state, and country. These enumeration areas are often used for comparison of standardized cancer rates in epidemiological investigations. Cumulatively, the proposed 15 AA capture or enable estimation of uncertainty in place of residence.

All attribute sets that enable or modify spatiotemporal relationships in cancer surveillance data exist as attribute associations with 2 exceptions. A cancer case has tumor and date of diagnosis among its primary key components. There is a subset of cases with information about the tumor coming from either the patient's recollection or patient death record, which constitutes approximately <2% of all cases as estimated for North Carolina. For this subset of tumor diagnoses, certainty in tumor and date of diagnosis depends on the certainty of patient identity. For the remaining >98% of cases, certainty in tumor and date of diagnosis is provided by biopsy, imaging, or clinical diagnosis, thus does not depend on a positively identified patient by a disease registry. Effectively, tumor and date of diagnosis represent attribute associations, which propagate uncertainty in residential geolocation for <2% of cases. For the remaining >98%, these are CCR-constrained sets,

meaning that for most cases, tumor (via date of diagnosis) does not propagate error probability into other AAs.

Compared to a more generic concept of the (perceived) infinity of space, residential space is better characterized by its finitude. At its most generic conception,²³ *place of residence* admits of all locations of (routine) human activity and its universe is approximated locally by emergency dispatch address points. By this convention, residential space extends to all areas except those that cannot support sustained human residences, such as within polygonal transportation routes or water bodies. That makes those areas useful for the placement of geocodes whose accuracy is uncertain, thus avoiding the assumptions of false precision. In a geographic coordinate system, the addressed area is discrete and uniquely described by X and Y coordinates. A residential geocode presents the association of a single set of coordinates with a place of residence, whose area it is proxy to. Stemming from statutory stewardship of E911 administrators, patient place of residence begins with a geocoded E911 address point or its proxies and extends to other residential features—parcels and digital building footprints—based on either the determination of its (proximity-based) record linkage relationship to them, or linkage enabled by site addresses that are attributes of digital parcels or building footprints (BFP).

In Table 1, we cite prior standards for attributes included in the proposed attribute sets. These existing standards are currently articulated by standard-setting organizations for cancer abstraction, and the placement of E911 emergency dispatch address points with the detailed information provided by Thornton et al 2020 and NENA 2015.^{12,24} The existing standards effectively provide building blocks of consensus about the granularity of features that we leverage in designating the 15 attribute sets.

Discussion

Novelty of the Proposed Approach

A review of the literature indicates that researchers have used error probability to assess the uncertainty of AAs—or tuple uncertainty—without necessarily referencing multiple stewardship.^{14,16-18} By contrast, our use of the term is specific to attribute sets stemming from conditionally privileged or statutory data stewards in a data stream. Approximately 91% of fields in the consolidated tumor record are constrained sets, such that cancer incidence data standard setting organizations or their proxies (CCRs) directly control single and cross-field domain constraints via statutory stewardship. The remaining 9% include patient demographic fields that collectively enable or modify spatiotemporal relationships based on the patient address at diagnosis,⁶ thereby significantly impacting confidence in published rates of cancer incidence. Understanding AAs in data thus becomes essential to managing the public expectations of the capabilities of cancer incidence data.

Current Limitations of Residential Geolocation Data at the Time of Diagnosis in Reference to Cancer Latency

Prior research has indicated a need for assessing

uncertainty in linking the environmental hazard testing data with cancer incidence and mortality data.²⁵ The presented methodological framework begins to address this concern by quantifying confidence in residential geolocation using AAs that appear in both the consolidated tumor record and environmental hazard testing (Table 1). While the public request to establish environmental origins of cancer clusters is strong, expanding the AA list to accommodate a connection between cancer incidence and specific environmental exposures is problematic. Given the current state of cancer surveillance data, information on residential exposure is limited to the time of diagnosis. However, the long latency period of developing cancer requires examination of the ties between residential history and environmental exposures, addressing periods of carcinogenesis or mutagenesis.^{26,27} Thus, incorporating latency of cancer development into AAs would require additional record linkage based on patient identity and address changes over time.²⁸ Unfortunately, for most CCRs, this scope of work is not funded.

Cancer Surveillance and Communication of Quantified Data Uncertainty

Recently, error probability has come into a greater focus in epidemiological investigations,²⁹⁻³³ as well as for risk-management decision-making.²⁹ We propose the most feasible way to manage expectations in the proactive scanning for cancer clusters by summarizing AA-specific error probability. This action presents a tangible step towards much-needed communication of quantified certainty in data. In communicating prospective scanning results, CCRs currently do not have the option of explicitly presenting geocoding uncertainty originating from multiple stewardship. Consequently, CCRs who are held culpable for all uncertainty arising from multiple stewardship in the data they curate must dispute narratives of uncertainty in their data in mass communications such as radio, television, and social media. Lack of explicit understanding of how much multiple data stewardship impacts the results of CCRs work diverts much of their resources to public discussions in which uncertainty in data plays a central role, and that comes at the expense of other work related to cluster investigations. In contrast, explicitly communicated quantified uncertainty (error probability) can help CCRs present to the public expected limitations of the existing cancer surveillance system.

In general, it is possible for at least some CCRs to minimize or mitigate the impact of uncertainty in residential geolocation on the outcomes of so-called high-profile cluster investigations by interactive (manual, case-by-case) geocoding and address research. This strategy may not be scalable to all cases for all CCRs, and therefore may not be available for proactive scanning. Considering that most CCRs are obligated to disclose any clusters that they detect, it seems apparent that a strategy to communicate risk strongly tied to summarized and quantified uncertainty in data is needed, and that capturing error probability as we have proposed is essential.

Summary

Ensuring a minimum level of comparability between scanned clusters—subsets of cancer cases—requires the assessment of confidence in attribute associations in the consolidated tumor record. Most of the uncertainty in residential geolocation of cancer surveillance data arises from the data provided by patients or their medical provider proxies instead of stewards of geographic reference data. Regardless, CCRs need a way to communicate quantified confidence in data whose constraints they do not control. Error probability—summarized over a group of cases—provides means to safely communicate these constraints without disclosing personal identifying information. If CCRs were primary stewards of all 15 AAs (Table1), then they would naturally employ cross-field and single-field domain constraints to provide confidence that 100% of error was eliminated. However, not being primary stewards of each of the 15 AAs, CCRs have a limited capacity to verify the entities. Thus, for CCRs, an AA presents a practical level of responsibility for assessing confidence in data constrained by limited validation capacity relative to primary stewardship.

For quantification of confidence in residential geolocation, much depends on how AAs are chosen. We have tried to keep our revised list of AAs minimal, including only those currently present in consolidated tumor records that can be justified as impactful to entity verification or subsequent analytic capability (Table1). Capturing confidence in residential geolocation at the attribute association level is crucial for determining comparability among clusters of cancer cases. In choosing attribute sets to account for error probability, we aimed to strike a balance between the granularity of residential geolocation and the replicability of metadata across CCRs. To facilitate the forging of consensus around the 15 AAs, we focus on AAs that enable record linkage to other data sets (eg, environmental hazard testing data), thereby establishing a more complete picture of analytic capacity. These include patient- and geography-based social determinants of health, and patient or patient profile-based consumer data used prospectively or retrospectively with analytics to assess past associations and predict outcomes. Our future research will focus on the methodology of how error probability in residential geolocation can be summarized.

Authors' Contributions

CK conceived this research, designed and coordinated the project, and drafted the original article and revisions. DI, CC, SB and KH reviewed, edited, and restructured the original article and revisions. All authors read and approved the final article.

Acknowledgements

CK wishes to acknowledge Hope Morgan and John Lay of the North Carolina Department of Public Safety for sharing statewide digital building footprints current as of 2016. The authors wish to acknowledge Nichole Baker and Robert Aldrich of the North Carolina CCR for their advice on CCR case consolidation, and other CCR-related topics.

References

1. Goodman M, Naiman JS, Goodman D, LaKind JS. Cancer clusters in the USA: What do the last twenty years of state and federal investigations tell us? *CRC Crit Rev Toxicol*. 2012;42(6): 474-490. doi:10.3109/10408444.2012.675315
2. Goodman M, LaKind JS, Fagliano JA, et al. Cancer cluster investigations: review of the past and proposals for the future. *Int J Environ Res Public Health*. 2014;11(2):1479-1499. doi:10.3390/ijerph110201479
3. Lemke D, Mattauch V, Heidinger O, Pedesma E, Hense HW. Detecting cancer clusters in a regional population with local cluster tests and Bayesian smoothing methods: a simulation study. *Int J Health Geogr*. 2013;12:54 doi:10.1186/1476-072X-12-54
4. Chen J, Roth RE, Naito AT, et al. Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of U.S. cervical cancer mortality. *Int J Health Geogr*. 2008;7:57. doi:10.1186/1476-072X-7-57
5. Jacquez GM. A research agenda: does geocoding positional error matter in health GIS studies? *Spat Spatiotemporal Epidemiol*. 2012;3(1):7-16. doi:10.1016/j.sste.2012.02.002
6. Klaus CA, Carrasco LE, Goldberg DW, Henry KA, Sherman RL. Use of attribute association error probability estimates to evaluate quality of medical record geocodes. *Int J Health Geogr*. 2015;14:26. doi:10.1186/s12942-015-0019-3
7. Cancer cluster guidelines update. Centers for Disease Control and Prevention website. Reviewed May 13, 2019. Accessed August 18, 2020. <https://www.cdc.gov/nceh/clusters/default.htm>
8. Waller LA, Gotway CA. *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons Inc; 2004.
9. Werner AK, Strosnider H, Kassinger C, Shin M; Sub-County Data Pilot Project Workgroup. Lessons learned from the Environmental Public Health Tracking Sub-County Data Pilot Project. *J Public Health Manag Prac*. 2018;24(5):E20-E27. doi:10.1097/PHH.0000000000000686
10. Hofferkamp J, ed. *Standards for Cancer Registries, Standard Data Edits*. Vol. IV. North American Association of Central Cancer Registries; 2020. <https://www.naaccr.org/standard-data-edits/>
11. Batini C, Scannapieco M. *Data Quality: Concepts, Methodologies and Techniques. Data-Centric Systems and Applications*. Springer; 2004.
12. Thornton ML, ed. *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary, Version 21*. 22nd ed. North American Association of Central Cancer Registries; 2020. Accessed August 18, 2020. <http://datadictionary.naaccr.org/default.aspx?c=10&Version=21>
13. O'Connor N. Reforming the U.S. approach to data protection and privacy. Council on Foreign Relations website. Published January 30, 2018. Accessed April 3, 2020. www.cfr.org/report/reforming-us-approach-data-protection
14. Sen P, Deshpande A, Getoor L. Representing tuple and attribute uncertainty in probabilistic databases. In: *Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007)*. IEEE; 2007:507-512. doi:10.1109/ICDMW.2007.11
15. Dalvi N, Suciu D. Efficient query evaluation on probabilistic databases. *Proceedings of International Conference on Very Large Databases (VLDB)*. Morgan Kaufmann Publishers; 2004:864-875. <http://www.vldb.org/conf/2004/RS22P1.PDF>
16. Singh S, Mayfield C, Shah R, et al. Database support for probabilistic attributes and tuples. *2008 IEEE 24th International Conference on Data Engineering*. 2008;1053-1061. doi:10.1109/ICDE.2008.4497514
17. Altenkrueger DE. A mathematical data/knowledge model: the formal association inference model FAIM. *Appl Math Model*. 1987;8:98-101.
18. Kon HB, Madnick SE, Siegel MD. *Good Answers from Bad Data: A Data Management Strategy*. MIT Sloan School of Management; 1995. https://www.researchgate.net/publication/5176173_Good_Answers_from_Bad_Data_a_Data_Management_Strategy
19. Goldberg DW, Wilson JP, Knoblock CA, Ritz B, Cockburn MG. An effective and efficient approach for manually improving geocoded data. *Int J Health Geogr*. 2008;7:60. doi:10.1186/1476-072X-7-60
20. North Carolina Center for Geographic and Information Analysis (2014). NC Master Address Dataset 2014 [database]. Accessed June 5, 2020. <https://www.nconemap.gov/datasets/nc-master-address-dataset-2014?showData=true>

21. North Carolina Center for Geographic and Information Analysis (2020). Parcels for North Carolina [database]. Accessed June 5, 2020. <https://www.nconemap.gov/pages/parcels>
22. North Carolina Department of Public Safety (2016). Statewide Building Footprints [database].
23. Xu D, Cui P, Zhu W, Yang S. Find you from your friends: graph-based residence location prediction for social media users. *IEEE Multimedia*. 2014;21:4:76-83.
24. National Emergency Number Association (NENA) Core Services Committee, NG Data Management Sub-Committee, and Site/Structure Address Points (SSAP) Workgroup. *NENA Information Document for Development of Site/Structure Address Point GIS Data for 9-1-1*. National Emergency Number Association; 2015. https://cdn.ymaws.com/www.nena.org/resource/resmgr/Standards/NENA-INF-014.1-2015_SSAP_INF.pdf
25. Hertz-Picciotto I. Toward a coordinated system for the surveillance of environmental health hazards. *Am J Public Health*. 1996;86(5): 638-641. doi:10.2105/ajph.86.5.638
26. Jacquez GM, Meliker J, Kaufmann A. In search of induction and latency periods: space-time interaction accounting for residential mobility, risk factors and covariates. *Int J Health Geogr*. 2007;6:35. doi:10.1186/1476-072X-6-35
27. Jacquez GM. Modeling carcinogenesis and cancer stages for latency estimation in dynamic geographic systems: linking genomic, cellular, individual and population levels. Presented at: Meetings of the North American Association of Central Cancer Registries; June 12, 2013; Austin, TX.
28. Wiese D, Stroup A, Maiti A, et al. Socioeconomic disparities in colon cancer survival: revisiting neighborhood poverty using residential histories. *Epidemiol*. 2020;31(5):728-735. doi:10.1097/EDE.0000000000001216
29. Burns CJ, Wright JM, Pierson JB, et al. Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments. *Environ Health Perspect*. 2014;122:1160-1165. doi:10.1289/ehp.1308062
30. Gilbert JA, Meyers LA, Galvani AP, Townsend JP. Probabilistic uncertainty analysis of epidemiological modeling to guide public health intervention policy. *Epidemics*. 2014;6:37-45. doi:10.1016/j.epidem.2013.11.002
31. Briggs DJ, Sabel CE, Lee K. Uncertainty in epidemiology and health risk and impact assessment. *Environ Geochem Health*. 2009;31:189-203.
32. Maziak W. Is uncertainty in complex disease epidemiology resolvable? *Emerg Themes Epidemiol*. 2015;12:7. doi:10.1186/s12982-015-0028-5
33. Sharif B, Kopec J, Wong H, et al. Uncertainty analysis in population-based disease microsimulation models. *Epidemiol Res Int*. 2012;122-136. doi:10.1155/2012/610405

Information Overload and What You Can Do About It

Michele Webb, CTR

Technology has changed how we process information and the amount of information we consume every day. Society has set expectations, too, and—whether realistic or not—may imply that if we do not keep up, we are not doing our jobs well.

The average American spends about 7 hours and 50 minutes a day processing digital media. In the next year, this will increase to 8 or more hours a day. In 2020, the rate of digital consumption rose, largely due to the pandemic, combined with the usual digital and traditional media such as television, radio, newspapers, and magazines.¹

Even if you are not a news junkie, we are bombarded daily with increasing amounts of new and repetitive information to the point that our brains are exhausted from constant processing. This response is called *information overload* and it happens when we attempt to process too much information, exceeding the brain's processing capacity. How each of us responds to information overload is usually based on our habits, lifestyles, and coping skills.

As cancer registrars, our exposure to digital media may even be higher. Sitting in front of multiple computer screens, gathering information from multiple electronic medical records, researching Web-based standards and coding manuals, deciphering edits, keeping up with software upgrades, understanding which standards were applicable in one year but not another—these are all part of our usual routine. Then, to top it off, staying on top of email, voice messaging, and department and facility updates also demand our attention. When you finally head home, there are cell phones, tablets, family obligations, news, social media, and the list goes on. No wonder we are on overload!

While we uniquely process information, the mental and physical impact of overload on our bodies is similar. It manifests itself much like a technical “brain freeze,” decreasing our ability to think clearly or efficiently, impairing decision-making, and decreasing productivity. The physical signs of information overload include increased blood pressure, headaches, irritability, insomnia, chronic fatigue, exhaustion, impatience, and intolerance.

To learn how to manage overload you need to understand its origins. There are 3 means by which we are exposed to information: personal, social, and business:

1. Personal. Here is an example of personal information exposure. Let's say you want to buy a new book to take on vacation. You flip open your laptop and navigate to your favorite bookseller's website. Almost immediately, hundreds of genres, titles, descriptions, buyer ratings, formatting options, and more are available. After you filter your list, you still have hundreds of books to choose from. You scroll down the first few pages, become overwhelmed, and either randomly pick a title or you shut everything down in frustration. But it does not stop there. Family, friends, community, and self-care also demand our attention.

2. Social. Scientists tell us that we accumulate about 2.5 quintillion bytes of information every day.² If you watched consecutive days of media reports about the pandemic last year, you already know how it feels to process large amounts of confusing or politically charged information. Social media, political and civic unrest, population health, and general intolerance for others has added complex layers to the information we consume on our computers, cell phones, tablets, television, and radio.

3. Business. Finally, a topic we are more comfortable talking about. But let's be realistic. The ever-changing reporting standards, employer's expectations of productivity, quality measures, performance rates, metrics, software systems, upgrades, reporting deadlines, and staffing shortages become part of the blur that assaults our brains every moment of the day.

Now that we know what information overload is and what it looks like, what can we do about it? How do we reduce or eliminate it? Can we manage it? Breaking bad habits or cycles that developed from not taking care of ourselves is difficult, but it can be done. Here are some things you can start doing right now:

- **Don't start the day with electronics!** Use the early part of the day to reflect and plan. Protect your energy and let yourself be fully present in the moment. If information overload is ramping up your blood pressure, step back or limit your intake by setting up specific or limited periods of time to take in information.
- **Stop multitasking.** The ability to multitask is a myth and reduces your productivity by as much as 40%. Prioritize your work and begin on the most important



tasks one at a time, followed by those that are least important. At the beginning of the day, take 3 to 5 minutes to plan your work. Identify the single most important task that must be completed that day, then get it done.

- **Disable electronic alerts, notifications, and devices.** This includes your laptops, desktop computers, cell phones, tablets, smart devices, email, or any other device that is disruptive. Let nonurgent phone calls go to voicemail. Set up filtering options in your email to prioritize incoming mail. Schedule a routine time each morning and afternoon to check email, return phone calls, or check in with your coworkers.
- **“Eat that Frog.”** Brian Tracy made this productivity hack famous when he introduced the concept of making your least favorite task of the day the one that is finished first, one “bite” at a time. Bring it to the top of our list, break it down into manageable chunks, and then finish it, one piece at a time. The more you procrastinate, the longer you are waiting to “eat the frog” and the bigger the consequences.
- **Follow the 2-minute rule.** If you have a lot of little tasks to do that will take under 2 minutes each, clump them together and knock them all out in a short amount of time. Stay focused and mindful of starting and completing each small task.
- **Practice gratitude, meditate, and enjoy life!** Connect with someone you care about each day. Take a walk outdoors or go to the gym. Spend time with your family, friends, favorite craft, or hobby. Practice gratitude and enjoy life.

Creating work-life balance depends on your ability to manage information. Develop an awareness of how your body reacts to overload and establish the habit of minimizing your exposure to media and sensationalized information. Plan and prioritize your day-to-day tasks. Turn off or remove distractions that drain your energy or focus. And, each day, find time to step away to turn off the electronics and enjoy family, friends, and places that recharge and refocus your energy.

References

1. Watson A. Time spent with digital vs. traditional media in the U.S. 2011-2022. Statista website. Published March 1, 2021. Accessed May 15, 2021. <https://www.statista.com/statistics/565628/time-spent-digital-traditional-media-usa/>
2. Huettich J. The ultimate guide to managing information overload. MindManager Blog website. Published February 10, 2020. Accessed May 15, 2021. <https://blog.mindmanager.com/blog/2020/02/202002the-ultimate-guide-to-managing-information-overload/>

Michele is a certified cancer registrar, speaker, author, and associate of SCL Healthcare in Colorado and Montana. She is committed to helping others grow and expand their influence as oncology health care leaders and mentors. Outside of work, Michele enjoys diamond painting, getting lost in a good book, and hanging out with her fur-babies, Dolly and Cooper. Your comments are welcome by email at michele@michelewebb.com.

Journal of Registry Management Continuing Education Quiz—SPRING 2021

LATE-STAGE DIAGNOSIS AND COST OF COLORECTAL CANCER TREATMENT IN TWO STATE MEDICAID PROGRAMS

After reading the article and completing the quiz, the participants will be able to:

- Describe the difference between the costs associated with late-stage and early-stage colon cancer diagnoses.
- Describe how timely and continuous Medicaid enrollment may lead to earlier stage at diagnosis, reduced costs, and improved outcomes.
- Understand the factors contributing to the higher costs of colon cancer in respect to the timing of Medicaid enrollment.

1. What is the acronym *CRC* in this article?
 - a) Cancer reporting center
 - b) Colorectal cancer
 - c) Central repository center
 - d) Centralized registry—colon
2. Which group of individuals were excluded from the analytic sample?
 - a) Individuals aged 21–64 years
 - b) Medicaid beneficiaries who enrolled 3 months or more after diagnosis
 - c) Medicaid beneficiaries who enrolled 12 months or more before diagnosis
 - d) Medicaid beneficiaries who enrolled 3 months or more prior to diagnosis
3. What was considered *late-stage* for this study?
 - a) American Joint Committee on Cancer clinical stage group 4
 - b) Localized, regional, or distant disease by Surveillance, Epidemiology, and End Results (SEER) Summary Stage
 - c) Regional, distant, or unknown disease by SEER Summary Stage
 - d) Regional or distant disease by SEER Summary Stage
4. According to Table 1, in California, what percentage of patients enrolled in Medicaid prior to diagnosis died within the first 6 months of diagnosis compared with patients enrolled in Medicaid after diagnosis?
 - a) 13.8% enrolled prior to diagnosis, 19.4% enrolled after diagnosis
 - b) 6.8% enrolled prior to diagnosis, 11.6% enrolled after diagnosis
 - c) 24.7% enrolled prior to diagnosis, 54.5% enrolled after diagnosis
 - d) 12.3% enrolled prior to diagnosis, 8.7% enrolled after diagnosis
5. Per Table 2, in Texas, patients diagnosed between 2000–2009 had what odds ratio of being diagnosed at a late stage if enrolled after diagnosis compared to those enrolled prior to diagnosis?
 - a) 1.54
 - b) 5.50
 - c) 3.96
 - d) 3.67
6. In Figure 1 for both Texas and California, the cost of hospital stays for distant disease showed what trend?
 - a) Costs increased for distant disease compared to localized disease.
 - b) Costs decreased for distant disease compared to localized disease.
 - c) There is not a difference in costs between distant and localized disease.
 - d) Costs could not be calculated for distant disease.
7. Which statement accurately describes how the timing of Medicaid enrollment affects stage at CRC diagnosis for this study?
 - a) Beneficiaries who enrolled in Medicaid prior to diagnosis had only early-stage diagnoses.
 - b) Beneficiaries who enrolled in Medicaid prior to diagnosis showed no difference in stage at diagnosis when compared to those enrolled after diagnosis.
 - c) Beneficiaries who enrolled in Medicaid prior to diagnosis were more likely to be diagnosed at a later stage of CRC.
 - d) Beneficiaries who enrolled in Medicaid prior to diagnosis were more likely to be diagnosed at an earlier stage of CRC.
8. Which costs were not included in this study?
 - a) Prescription costs
 - b) Hospital stays
 - c) Ambulatory costs
 - d) End of life costs
9. In 2010, the cost of CRC care was second highest, second only to which other type of cancer?
 - a) Pancreatic
 - b) Lymphoma
 - c) Breast
 - d) Lung
10. According to the study, which of the following statements is true?
 - a) Timely and continuous Medicaid enrollment for those eligible may lead to earlier stage at diagnosis, reduced costs, and improved outcomes.
 - b) The study proved there was no benefit to timely and continuous Medicaid enrollment.
 - c) There was no reduction in costs or improved outcomes based on timing of Medicaid enrollment.
 - d) Better assessment of those patients not eligible for Medicaid may lead to improved outcomes.

Purchase Quiz to Earn CE:

1. Go to <http://www.cancerregistryeducation.org/jrm-quizzes>
2. Select quiz and “Add to Cart” (You may be prompted to login using your NCRA login).
3. Continue through the checkout process.
4. Once purchase is complete, the quiz will load automatically into “My Learning Activities” page.



Oncology Data Management Software Revolutionized That is the KACI® Difference

KACI (Knowledge Assistant for Clinical Information) The New Platform with Breakthrough Technology

KACI readily provides insights to support and grow your organization.

- ☒ Leverage existing registry practices to quickly and easily create dashboard graphs, reports, and information to facilitate:
 - Accreditation compliance
 - Analysis of disease patterns impacting the community
 - Trends in service utilization for staffing and capital purchases
- ☒ Improve abstracting performance with KACI's intuitive workflow, automated standards and guidelines, and benchmarking capabilities.



HL7 FHIR R4 Enabled • Scalable Modern Architecture • SOC 2 Type II Audited

Schedule a Demo

neuralframe.com/schedule-a-demo

Give us a call: 855.494.5900

About NeuralFrame

Founded in 2017, NeuralFrame has brought together an expert team of clinicians, systems engineers, CTRs, and executives passionate about generating better medical data for better patient outcomes.

©2021 NeuralFrame, Inc. All Rights Reserved

National Cancer Registrars Association

CALL FOR PAPERS

Danette A. Clark, BS, RMA, AAS, CTR | EDITOR-IN-CHIEF, *JRM*

The *Journal of Registry Management*, official journal of the National Cancer Registrars Association (NCRA), announces a call for original manuscripts on registry methodology or research findings related to the 7 subjects listed below and related topics.

Topics:

1. **Birth Defects Registries**
2. **Cancer Registries**
 - a. **AJCC TNM Stage**
 - b. **Cancer and Socioeconomic Status**
 - c. **Cancer and Health Disparities**
3. **Trauma Registries**
4. **Recruitment, Training, and Retention**
5. **Public Relations**
6. **Quality Review**
7. **Registry Management**

Contributed manuscripts are peer-reviewed prior to publication. Manuscripts of the following types may be submitted for publication:

1. **Methodology Articles** addressing topics of broad interest and appeal to the readership, including methodological aspects of registry organization and operation.
2. **Research articles** reporting findings of original, reviewed, data-based research.
3. **Primers** providing basic and comprehensive tutorials on relevant subjects.
4. **“How I Do It” Articles** describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. The *“How I Do It”* feature in the *Journal* provides registrars with an informal forum for sharing strategies with colleagues in all types of registries.
5. **Opinion papers/editorials** including position papers, commentaries, essays, and interviews that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management.
6. **Bibliographies** which are specifically targeted and of significant interest will be considered.
7. **Letters to the Editor** are also invited.

Address all manuscripts to: Danette Clark, BS, RMA, AAS, CTR, Editor-in-Chief, *Journal of Registry Management*, (973) 971-5189, JRMEditor@ncra-usa.org.

Manuscript submission requirements are given in “Information for Authors” found near the back of each *Journal* and on the NCRA website at <http://www.ncra-usa.org/jrm>.

Journal of Registry Management

INFORMATION FOR AUTHORS

Journal of Registry Management (JRM), the official journal of the National Cancer Registrars Association, invites submission of original manuscripts on topics related to management of disease registries and the collection, management, and use of cancer, trauma, AIDS, and other disease registry data. Reprinting of previously published material will be considered for publication only when it is of special and immediate interest to the readership. *JRM* encourages authorship by Certified Tumor Registrars (CTRs); special value is placed on manuscripts with CTR collaboration and publication of articles or texts related to the registry profession. CTR continuing education (CE) credits are awarded; a published chapter or full textbook article equals 5 CE hours. Other published articles or documents equal CE hours. All correspondence and manuscripts should be addressed to the Danette Clark, BS, RMA, AAS, CTR, Editor-in-Chief at JRMEditor@ncra-usa.org or (973) 971-5189.

Manuscripts may be submitted for publication in the following categories: **Articles** addressing topics of broad interest and appeal to the readership, including **Methodology papers** about registry organization and operation; **Research papers** reporting findings of original, reviewed, data-based research; **Primers** providing tutorials on relevant subjects; and **"How I Do It"** papers are also solicited. **Opinion papers/editorials** including position papers, commentaries, and essays that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management; **Letters to the Editor**; and specifically-targeted **Bibliographies** of significant interest are invited.

The following guidelines are provided to assist prospective authors in preparing manuscripts for the *Journal*, and to facilitate technical processing of submissions. Failure to follow the guidelines may delay consideration of your manuscript. Authors who are unfamiliar with preparation and submission of manuscripts for publication are encouraged to contact the Editor for clarification or additional assistance.

Submission Requirements

Manuscripts. The terms manuscripts, articles, and papers are used synonymously herein. Email only submission of manuscripts is encouraged. If not feasible, submit the original manuscript and 4 copies to the Editor. Manuscripts should be double-spaced on white 8-1/2" x 11" paper, with margins of at least 1 inch. Use only letter-quality printers; poor quality copies will not be considered. Number the manuscript pages consecutively with the (first) title page as page one, followed by the abstract, text, references, and visuals. The accompanying cover letter should include the name, mailing address, email address, and telephone number of the corresponding author. For electronic submission, files should be IBM-compatible format in Corel WordPerfect™, Microsoft® Word for Windows®, or converted to ASCII code.

Manuscripts (Research Articles). Articles should follow the standard format for research reporting (Introduction, Methods, Results, Discussion, References), and the submission instructions outlined above. The introduction will normally include background information, and a rationale/justification as to why the subject matter is of interest. The discussion often includes a conclusion subsection. Comprehensive references are encouraged, as are an appropriate combination of tables and figures (graphs).

Manuscripts (Methodology/Process Papers). Methodology papers should follow the standard format for research reporting (Introduction, Methods, Results, Discussion), or for explanatory papers not reporting results (Introduction, Methods, Discussion), as well as the submission instructions outlined above.

Manuscripts ("How I Do It" articles). The "How I Do It" feature in the *Journal* provides registrars with a forum for sharing strategies with colleagues in all types of registries. These articles describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. When shared, these innovations can help registry professionals improve their skills, enhance registry operations, or increase efficiency.

"How I Do It" articles should be 1,500 words or less (excepting references) and can contain up to 2 tables or figures. To the extent possible, the standard headings (Introduction, Methods, Results, Discussion) should be used. If results are not presented, that section may be omitted. Authors should describe the problem or issue, their solution, advantages (and disadvantages) to the suggested approach, and their conclusion. All submitted "How I Do It" articles will have the benefit of peer/editorial review.

Authors. Each author's name, degrees, certifications, title, professional affiliation, and email address must be noted on the title page exactly as it is to appear in publication. The corresponding author should be noted, with mailing address included. Joint authors should be listed in the order of their contribution to the work. Generally, a maximum of 6 authors for each article will be listed.

Title. Authors are urged to choose a title that accurately and concisely describes the content of the manuscript. Every effort will be made to use the title as submitted; however, *Journal of Registry Management* reserves the right to select a title that is consistent with editorial and production requirements.

Abstract. A brief abstract must accompany each article or research paper. The abstract should summarize the main point(s) and quickly give the reader an understanding of the manuscript's content. It should be placed on a page by itself, immediately following the title page.

Length. Authors are invited to contact the Editor regarding submission of markedly longer manuscripts.

Style. Prepare manuscripts using the *American Medical Association Manual of Style*, 11th ed. (2020).

Visuals. Use visuals selectively to supplement the text. Visual elements—charts, graphs, tables, diagrams, and figures—will be reproduced exactly as received. Copies must be clear and properly identified, and preferably emailed. Each visual must have a brief, self-explanatory title. Submit each visual on a separately numbered page at the end of the manuscript, following the references.

Attribution. Authors are to provide appropriate acknowledgment of products, activities, and support especially for those articles based on, or utilizing, registry data (including acknowledgment of hospital and central registrars). Appropriate attribution is also to be provided to acknowledge federal funding sources of registries from which the data are obtained.

References. References should be carefully selected, and relevant. References must be numbered in order of their appearance in the text. At the end of the manuscript, list the references as they are cited; do not list references alphabetically. Journal citations should include author, title, journal, year, volume, issue, and pages. Book citations should include author, title, city, publisher, year, and pages. *Authors are responsible for the accuracy of all references.* Examples:

1. LeMaster PL, Connell CM. Health education interventions among Native Americans: a review and analysis. *Health Education Quarterly*. 1995;21(4):521-538.
2. Hanks GE, Myers CE, Scardino PT. Cancer of the prostate. In: DeVita VT, Hellman S, Rosenberg SA. *Cancer: Principles and Practice of Oncology*. 4th ed. Philadelphia, PA: J.B. Lippincott Co.; 1993:1073-1113.

Key words. Authors are requested to provide up to 5, alphabetized key words or phrases which will be used in compiling the Annual Subject Index.

Affirmations

Copyright. Authors submitting a manuscript do so on the understanding that if it is accepted for publication, copyright in the article, including the right to reproduce the article in all forms and media, shall be assigned exclusively to NCRA. NCRA will not refuse any reasonable requests by the author(s) for permission to reproduce any of his or her contributions to the *Journal*. Further, the manuscript's accompanying cover letter, signed by all authors, must include the following statement: "We, the undersigned, transfer to the National Cancer Registrars Association, the copyright for this manuscript in the event that it is published in *Journal of Registry Management*." Failure to provide the statement will delay consideration of the manuscript. It is the author's responsibility to obtain necessary permission when using material (including graphs, charts, pictures, etc) that has appeared in other published works.

Originality. Articles are reviewed for publication assuming that they have not been accepted or published previously and are not under simultaneous consideration for publication elsewhere. If the article has been previously published or significantly distributed, this should be noted in the submission for consideration.

Editing

Journal of Registry Management reserves the right to edit all contributions for clarity and length. Minor changes (punctuation, spelling, grammar, syntax) will be made at the discretion of the editorial staff. Substantive changes will be verified with the author(s) prior to publication.

Peer Review

Contributed manuscripts are peer-reviewed prior to publication, generally by 3 reviewers. The *Journal* Editor makes the final decision regarding acceptance of manuscripts. Receipt of manuscripts will be acknowledged promptly, and corresponding authors will be advised of the status of their submission as soon as possible.

Reprints

Authors receive 5 complimentary copies of the *Journal* in which their manuscript appears. Additional copies of reprints may be purchased from the NCRA Executive Office.

Journal of Registry Management

NCRA Executive Office
1330 Braddock Place
#520
Alexandria, VA 22314



Printed on SFI fiber sourcing paper with non-petroleum,
vegetable based inks and manufactured with renewable electricity



PRESORTED
STANDARD
U.S. POSTAGE
PAID
DULLES, VA
PERMIT #6418

ADDRESS SERVICE REQUESTED